

RESEARCH ARTICLES

The role of visual attention in multiple object tracking: Evidence from ERPs

MATTHEW M. DORAN AND JAMES E. HOFFMAN
University of Delaware, Newark, Delaware

We examined the role of visual attention in the multiple object tracking (MOT) task by measuring the amplitude of the N1 component of the event-related potential to probe flashes presented on targets, distractors, or empty background areas. We found evidence that visual attention enhances targets and suppresses distractors (Experiments 1 and 3). However, we also found that when tracking load was light (two targets and two distractors), accurate tracking could be carried out without any apparent contribution from the visual attention system (Experiment 2). Our results suggest that attentional selection during MOT is flexibly determined by task demands, as well as tracking load, and that visual attention may not always be necessary for accurate tracking.

We often need to simultaneously monitor several objects in our environment. Everyday examples include team sports such as basketball, in which each player needs to keep track of players on his own team as well as his opponents, and automobile drivers, who need to know the positions and motion trajectories of cars in neighboring lanes while monitoring oncoming traffic. The relative ease with which we do this might imply that we simply monitor all the objects in our visual field, but, in fact, it is easy to show that this is not the case, as demonstrated in the multiple object tracking (MOT) task.

The MOT task was developed by Pylyshyn and Storm (1988) as a laboratory analogue of the divided visual attention situation described above. In this task, observers are first shown a small set (8–16) of identical objects (e.g., black circles). A subset of them are designated as targets by briefly flashing them, and then all the objects move randomly about the display while the observer attempts to keep track of the targets. At the end of the trial, the objects become stationary, and the observer selects the target objects. The main finding is that observers can accurately track approximately four objects and that once this limit is exceeded, accuracy declines precipitously. Thus, divided attention appears to be limited to four objects, a limit that also appears in a variety of other visual attention tasks (see Cowan, 2001), such as whole report (Sperling, 1960), change detection (Luck & Vogel, 1997; Pashler, 1988), and subitizing (Mandler & Shebo, 1982; Trick & Pylyshyn, 1994).

Here, we consider the question of how people selectively attend to targets and exclude distractors in the MOT task. Several different selection mechanisms are known

to operate in vision, including selection by features, locations, and objects. For example, when we search for a shape such as a letter, we *activate* the features of that shape (angular, horizontal line), with the result that only visual objects that possess those features interfere with our search (Bacon & Egeth, 1994). This mechanism does not seem to be at work in MOT, because all the objects are identical and cannot be distinguished on the basis of their features.

Selection by Location

Selection by location has been extensively studied, beginning with Helmholtz (1867/1925; see Wright & Ward, 2008, for a history of these early investigations of visual attention), who noted that we can voluntarily pay attention to peripheral locations without changing our fixation, an ability known as *covert orienting* (Posner, 1980). The result is faster and more accurate detection and identification of information appearing in the attended area (Averbach & Coriell, 1961; C. W. Eriksen & Hoffman, 1972, 1973; C. W. Eriksen & Yeh, 1985; Posner, 1980; Posner, Snyder, & Davidson, 1980). This mechanism has been likened to a *spotlight* illuminating the attended area. The spotlight of attention appears to have a limited spatial resolution (Intriligator & Cavanagh, 2001) and may include spatial locations adjacent to the attended area (C. W. Eriksen & Hoffman, 1972, 1973; Hoffman & Nelson, 1981). In addition, it has been likened to a zoom lens that can be focused on a small area to provide detailed information about a few objects or can have a broad focus to yield less detailed information about many objects (C. W. Eriksen & St. James, 1986; C. W. Erik-

sen & Yeh, 1985). The attentional spotlight can also be divided and simultaneously directed to different spatial locations (Awh & Pashler, 2000; Franconeri, Alvarez, & Enns, 2007; Malinowski, Fuchs, & Müller, 2007; McMains & Somers, 2004; Müller, Malinowski, Gruber, & Hillyard, 2003).

Object-Based Attention

In addition to feature- and location-based selection mechanisms, researchers have identified a third possible selection mechanism based on objects (see Scholl, 2001, for a review), although it is presently unclear whether this system might ultimately be reduced to selection based on some combination of features and location. For example, O'Craven, Downing, and Kanwisher (1999) presented observers with a partially transparent picture of a house superimposed on a similarly formatted picture of a face. On some trials, the face moved back and forth while the house remained stationary, and observers were required to monitor the direction of motion. O'Craven et al. found that areas of the brain that are known to be activated by faces (*the fusiform face area*) were also active here, whereas *house areas* (parahippocampal place area) were relatively silent. This pattern was reversed when the observers attended to the house image. These results show, first, that superimposed images can be separately attended, even though they occupy the same general location and, second, that attending to one feature of an object (its motion) results in the selection of other aspects of that same object (its shape: face or house). The selection process itself, which is responsible for admitting faces and excluding houses, might be based on the feature of motion. However, the unit of selection appears to be the whole object and not just individual features (Blaser, Pylyshyn, & Holcombe, 2000; Duncan, 1984; Egly, Driver, & Rafal, 1994).

Object Indexing

Pylyshyn and colleagues (Pylyshyn, 2001, 2003; Pylyshyn & Storm, 1988; Sears & Pylyshyn, 2000) have offered a theory of performance in the MOT task that rejects both features and spatial locations as the basis of selection. Pylyshyn (2003) acknowledged the need for a focal attention system that can select different locations in the visual world but argued that there is an earlier, more primitive selection mechanism that serves the function of *indexing* or pointing to objects. Importantly, he maintained that "objects are indexed directly, rather than via their properties or their locations" (p. 202). Each object to be tracked is assigned an index that passively adheres to that object as it moves. The finding that observers can track only four objects suggests a corresponding limit in the number of available indexes (Pylyshyn & Storm, 1988; Scholl, 2001; Scholl & Pylyshyn, 1999; for a review, see Scholl, 2009).

According to Pylyshyn (e.g., 2001, 2003), object tracking is carried out by an indexing system that is separate from the location-based spatial attention system. Indexes allow people to determine properties of tracked objects, such as their color, location, or shape, but these properties themselves are not the basis for tracking. If one needs to know the color or some other property of a tracked

object, *visual attention* can be allocated to a particular object, using the *index* as a pointer to the relevant object. Thus, it is the indexing system that distinguishes targets from distractors, and this mechanism is distinct from the purely location-based spotlight of attention. This is an important claim because it is pivotal to the assertion that the proposed indexing system, which is presumed to be the principal basis for attending to objects, is a separate selection mechanism and cannot simply be reduced to some combination of feature- and location-based attention systems. The initial evidence for this claim was provided by Pylyshyn and Storm (1988), who argued that a unitary focus of spatial attention could not be moved among tracked objects quickly enough to support the accurate performance that was observed in the typical MOT experiment. They argued instead for the existence of a set of indexes that can track the objects in parallel.

This argument, however, does not exclude a variation of attentional selection in which attention may be split into multiple, independent foci (Awh & Pashler, 2000; Franconeri et al., 2007; Malinowski et al., 2007; McMains & Somers, 2004; Müller et al., 2003), each devoted to one of the tracked objects (Cavanagh & Alvarez, 2005). In fact, the multifocal attention model appears quite similar to the indexing model, with indexes and attentional foci playing similar roles. Both models suggest that MOT is accomplished by virtue of a limited resource's (indexes or visual attention) being simultaneously allocated to the tracked objects.

Object Indexing Versus Visual Attention

Current evidence does not appear to discriminate between these two models. For example, a variety of evidence has accumulated supporting the *object-based nature of tracking*. When observers attempt to track one end of a barbell type of object, tracking accuracy declines because they cannot keep the two ends separate (Scholl, Pylyshyn, & Feldman, 2001), as would be expected if they were tracking a single object rather than part of an object. Similarly, observers have trouble tracking substances that flow smoothly from one location to another (vanMarle & Scholl, 2003), perhaps because the visual system maintains a distinction between objects and substances, and tracking can be applied only to the former. Although this evidence is consistent with the claim that at least sometimes the *unit* of selection is an object, it does not rule out the possibility that the *mechanism* of selection is still spatial in nature.

There is also evidence that seems to argue against a simple indexing mechanism, particularly the idea that limits on object tracking are due to an underlying limit on the number of indexes. First, the number of objects that can be accurately tracked is not fixed at four (or any other number, for that matter). Alvarez and Franconeri (2007) showed that this number could vary from one object to as many as eight, depending on factors such as speed and spatial separation between the objects. They suggested that the underlying limit in MOT was a finite processing capacity that is shared among the tracked objects. This system can track only a small number of items when con-

ditions are difficult and each object demands a large share of the available capacity, but it can track a much larger number of objects when tracking is easy and each object makes small demands on the total capacity.

The limited capacity mechanism identified by Alvarez and Franconeri (2007) might be visual attention, which becomes “diluted” as it is divided among additional locations containing target objects. And just as visual attention has a limited spatial resolution, so too does MOT (Franconeri, Lin, Pylyshyn, Fisher, & Enns, 2008; Shim, Alvarez, & Jiang, 2008). Several experiments have shown that one important source of errors in MOT is the likelihood that observers will mistakenly start tracking a distractor object if it passes too close to a target (O’Hearn, Landau, & Hoffman, 2005; Pylyshyn, 2004). This kind of error seems to follow naturally from the idea that the spotlight of visual attention is limited in its spatial resolving power, as has been noted by many other investigators (e.g., B. A. Eriksen & Eriksen, 1974; C. W. Eriksen & Hoffman, 1972, 1973; Hoffman & Nelson, 1981; Intriligator & vanaganah, 2001).

One way to evaluate the role of visual attention in MOT is to require observers to detect brief probes presented during the MOT task. If a separate focus of attention is devoted to each target, probes presented on targets should be detected better than probes presented on distractors or background areas. This prediction rests on the observation that when visual attention is allocated to a location, all information in the attended region receives enhanced processing, even when it is irrelevant (Luck, Fan, & Hillyard, 1993). Therefore, we would expect that if observers allocate visual attention to an object for the purpose of tracking it, there will also be benefits in processing other information, such as probe objects, that appear in the same spatial neighborhood as the tracked object. Note that, according to Pylyshyn’s (2003) conception of indexes, such a finding would not necessarily follow from indexing theory, because indexes allocated to tracked objects do not automatically provide access to any of those objects’ properties, let alone other visual information, such as probes, that are distinct objects in their own right. In addition, indexes do not provide explicit information about an object’s spatial location. Spatial location, like color, is simply another object property that can be retrieved, if needed, by using the index as a pointer, so there is no reason why probes that are *near* to tracked objects should enjoy any advantage over more distant ones.

Nevertheless, Pylyshyn (2006) performed just such a probe experiment and found that, surprisingly, probe detection was better for both target and background locations, as compared with distractor locations. This result does not easily fit with the idea that visual attention is being preferentially allocated to targets. Instead, Pylyshyn (2006) suggested that distractors are being inhibited. This makes sense, because it is distractors passing close to targets that pose the main threat to good performance in the MOT task. A more recent article (Pylyshyn, Haladjian, King, & Reilly, 2008) showed that distractors that could be preattentively segregated from targets by appearing in a different depth plane received less inhibition, reinforcing

the idea that inhibition is applied to distractors when they are close to and could be mistaken for targets.

Although probe detection experiments appear to provide intriguing evidence about how attention is allocated during MOT, we have to keep in mind that observers are being placed in a dual-task situation and they may be adopting strategies that are different from those used in the “pure” MOT task. In addition, better probe detection could reflect attentional effects at a variety of levels of processing (Vogel, Woodman, & Luck, 2005). In an alternative approach, the role of visual attention in MOT is investigated by examining brain activity, in the form of event-related brain potentials (ERPs), elicited by task-irrelevant *probe flashes* presented on targets, distractors, or background areas. One advantage of this approach is that ERP amplitude can provide a measure of attention without requiring observers to make any judgment about the probes and, therefore, may constitute a relatively unobtrusive measure of attention during MOT. In addition, different ERP components, such as the N1 and P300, appear to index different stages of attention, so such measures may also provide information about the level at which attention operates during MOT.

Extensive prior research shows that the amplitudes of two *early* ERP components (the P1 and N1) that are generated in the extrastriate visual cortex are larger for stimuli appearing in attended spatial locations (Di Russo, Martínez, & Hillyard, 2003; Di Russo, Martínez, Sereno, Pitzalis, & Hillyard, 2002; Hillyard & Anllo-Vento, 1998; Luck, Woodman, & Vogel, 2000; Mangun & Hillyard, 1988; Martínez et al., 2001). Importantly, P1 and N1 amplitudes have been shown to be unaffected by attention to stimulus features such as color or size (Hillyard & Anllo-Vento, 1998; Hillyard & Münte, 1984). More recent experiments have shown similar effects for probes appearing on attended surfaces or objects (Martínez, Ramanathan, Foxe, Javitt, & Hillyard, 2007; Martínez, Teder-Sälejärvi, & Hillyard, 2007; Martínez et al., 2006; Valdes-Sosa, Bobes, Rodriguez, & Pinilla, 1998), but these object-based effects could be mediated by visual attention’s conforming to the shapes of attended objects and surfaces. Current evidence, then, indicates that the amplitude of these two components is specific to visuospatial attention and is not affected by feature-based attention. If visual attention is preferentially allocated to targets during MOT, we would expect larger P1/N1 components for target flashes, relative to both distractor and background flashes. Alternatively, if visual attention suppresses distractor items during MOT, as was suggested by Pylyshyn (2006), we would expect to observe larger P1/N1 components for both target and background flashes, relative to distractor flashes.

Drew, McCollough, Horowitz, and Vogel (2009) applied this logic to probes appearing on targets, moving distractors, stationary distractors, and blank display areas. They found that target probes elicited larger P1 and N1 components than did other probes, suggesting that visual attention enhances tracked objects during MOT. The present Experiment 1 is similar in many respects to the study carried out by Drew et al. but uses moving texture-defined objects rather than the luminance-defined objects used by

Drew et al., in an effort to equate the physical properties of probes occurring on objects and background areas (see Flombaum & Scholl, 2009; Flombaum, Scholl, & Pylyshyn, 2008, who used similar stimuli in an MOT task). In addition, this experiment served as a comparison for later experiments that required responses to the probes (Experiment 2) and increased tracking load (Experiment 3).

EXPERIMENT 1 Irrelevant Probes

In this experiment, participants tracked two out of four objects while probe flashes were presented on targets, distractors, or the background. All the probe flashes were irrelevant to the observers, since they did not require any behavioral response and could not be used to aid tracking performance, since the probes were equally likely to appear in any of these locations. As a result, we were able to use the amplitude of the P1/N1 components generated by the probe flashes to measure the allocation of visual attention while the participants were performing the MOT task by itself.

Method

Participants. Eighteen right-handed, neurologically normal volunteers (ages, 18–30 years) participated in this experiment. With the exception of author M.M.D., all were naive as to the purpose of the experiment and were paid \$10/h for their participation. All the participants reported normal or corrected-to-normal acuity and provided informed consent.

Stimuli and Procedure. The stimuli were generated on a Dell 2.99-GHz computer running custom software written with Blitz3D (Sibly, 2005) and were presented on an 18-in. Dell CRT ($1,024 \times 768$ pixel resolution; 75-Hz frame rate). Eye fixation was monitored using a Tobii x50 50-Hz eyetracker (Tobii Technology, Stockholm, Sweden) controlled by a Sony 2.86-GHz computer. Testing was conducted in a dimly lit, electrically shielded room with a chinrest maintaining a 70-cm viewing distance. Under these conditions, the total viewable area of the screen subtended approximately $27.5^\circ \times 21.1^\circ$, but the stimuli appeared only in the center $19.7^\circ \times 14.0^\circ$ of the screen, with the remainder of the screen being black.

The participants tracked two out of four objects (squares, 1.43°) that moved haphazardly throughout the display. The objects and background were constructed from randomly generated black (0.10 cd/m^2) and white (76.75 cd/m^2) pixels, resulting in displays in which the objects were invisible when stationary (examples of these displays can be viewed at <http://hoffman.psych.udel.edu/TextureMOT/index.html>). The primary advantage of this display format is that objects and background are physically identical during tracking (Flombaum et al., 2008). At the beginning of each trial, the objects were outlined by black placeholders so that they were visible to the observer. The initial location of each object was selected randomly, with the constraint that objects could not overlap.

The participants initiated each trial by pressing a mouse button that caused two of the objects' placeholders to blink on and off five times for 200 msec, identifying those objects as targets. Once the cuing was complete, the placeholders remained invisible while the objects and background moved for a period of 10 sec. When the motion ended, the placeholders reappeared, and the participants used the mouse to select the targets. Immediate feedback was provided by changing the color of the placeholder to green for a correctly selected target and to red for an incorrectly chosen distractor. This selection phase ended once the participants selected two objects. Throughout the trial, a blue fixation circle (0.34°) was present at the center of the display, and the participants were instructed to maintain fixation. Additional feedback was provided at the end of the trial

about the percentage of correct tracking trials (i.e., the percentage of trials on which both targets were correctly chosen), as well as the number of eye movements and blinks that were detected during the trial.

Motion trajectories were computed offline for each participant. At the beginning of each trial, randomly selected speeds ranging between $\pm 2^\circ/\text{sec}$ in the horizontal and vertical directions were assigned to objects and the background texture. In order to keep the objects visible, these initial speeds were constrained so that the objects and background always began with nonzero velocities. For the same reason, the objects and background moved in different directions with different velocities throughout the trial. In order to produce a *wandering* motion, the horizontal/vertical speeds of each object and the background were adjusted randomly and independently in increments between $\pm 0.5^\circ/\text{sec}$, with the constraint that the speeds never exceeded $2^\circ/\text{sec}$. On average, the objects and background moved $1.2^\circ/\text{sec}$. The object motions were also constrained so that the objects reflected off of each other and the edges of the display.

In order to measure the allocation of attention during tracking using ERPs, 12 probe flashes (white squares that were identical in size to the objects; 76.75 cd/m^2) were presented (one at a time) during each MOT trial. The first probe was presented 400–600 msec after the initiation of motion on each trial, and the last probe occurred at least 900 msec before the motion ended. Probes remained visible for 106 msec (i.e., 8 refresh frames), and the time between probes was randomly drawn from the 400- to 600-msec interval (i.e., 30–45 refresh frames). On each trial, probes were distributed equally on targets, distractors, and the background in both visual fields and were presented in random order. All the probes moved along the same trajectory as the surface on which they occurred. In other words, object probes followed the same trajectories as the probed object, and background probes followed the same trajectory as the background. Since the background moved according to similar constraints as the objects and both were composed of random textures, the background and object probes were physically comparable. Probes that appeared on the background were presented at randomly chosen locations so that the centers of the probes were at least 2.86° from the edge of the display and from the centers of any objects. The participants were informed that the probes could appear on any of the objects or the background, that they could not be used as an aid in tracking, and that they were irrelevant to the tracking task.

Each participant viewed a total of 100 experimental MOT trials in a single session. As a result, the participants viewed a total of 200 probe presentations in each condition, created by crossing visual field (left vs. right) with probe location (target, distractor, or background). For each participant, pairs of trials were identical, except that the identities of the targets and distractors were switched. In other words, for each trial, there was a companion trial that had identical motion trajectories and probe presentations, but the targets on one trial were the distractors on a companion trial. In addition, the participants viewed approximately 10 randomly generated practice trials prior to the beginning of the experimental trials.

Electrophysiological recording. The electroencephalogram (EEG) was recorded with an Electrical Geodesics Inc. (EGI; Eugene, OR) 128-channel Geodesic Sensor Net with individual electrode impedances kept below $50\text{--}75\text{ k}\Omega$, as recommended by the manufacturer. The data were referenced online to the vertex, band-pass filtered from 0.01 to 80 Hz, and digitized at 200 Hz. Subsequent processing was performed offline using EGI Net Station 4.1.2. The data were band-pass filtered from 1 to 55 Hz and then were segmented using an epoch that began 100 msec prior to the onset of the probe flash and ended 800 msec after. Net Station's artifact detection routines were then applied. Individual channels were marked as bad if there was zero variance, the fast average amplitude exceeded $200\text{ }\mu\text{V}$, or the differential average amplitude exceeded $200\text{ }\mu\text{V}$. Individual segments were rejected if they contained eye movements or blinks (threshold = $70\text{ }\mu\text{V}$) or if more than 10 channels were marked as bad. For the remaining segments, bad channels were replaced by interpolating from surrounding channels. Finally, the segments

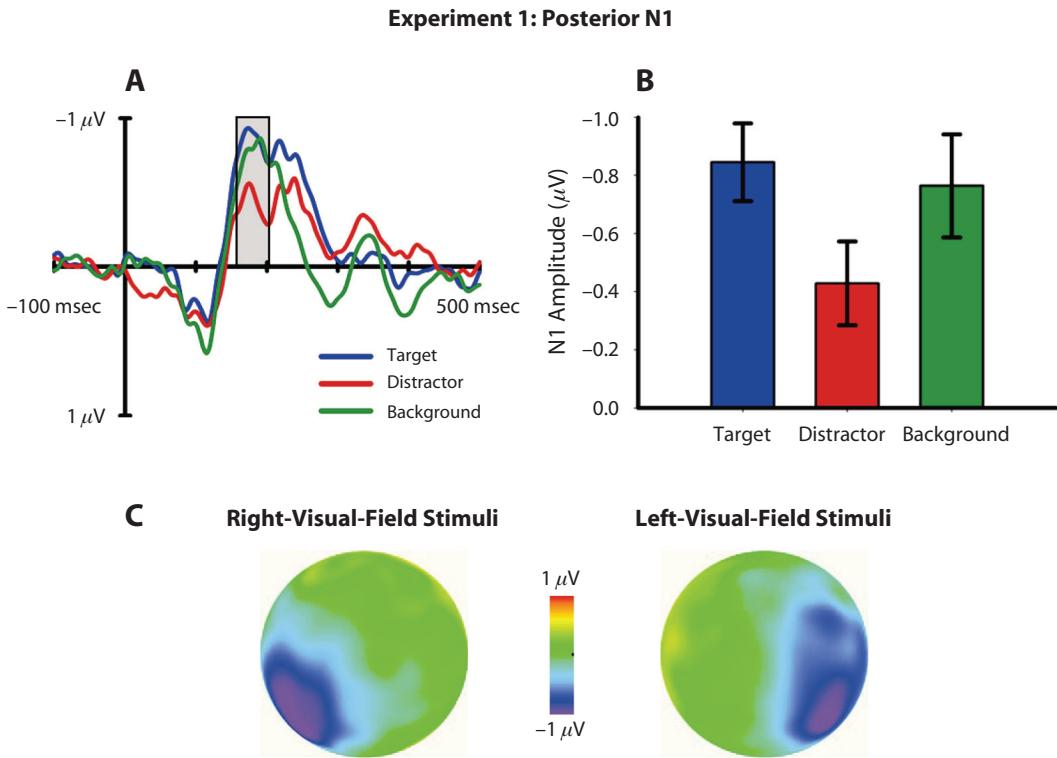


Figure 1. Posterior N1 results from Experiment 1. (A) Grand averaged ERP waveforms elicited by contralateral visual field stimuli. The shaded region corresponds to the analysis time window. (B) Mean N1 amplitude in the measurement time window. (C) Isopotential scalp maps for left- and right-visual-field target probes 175 msec poststimulus. Note that target and background probes produced larger amplitude N1 components than did distractor probes. Error bars are ± 1 standard error.

were averaged, rereferenced to the average reference, and baseline corrected using a 100-msec prestimulus interval.

Results

As was expected, tracking accuracy was fairly high. On average, the participants correctly selected both target objects on 92.6% of the trials ($SD = 4.6$). Using the formula provided by Pylyshyn and Annan (2006; see also Hulleman, 2005, for a similar formula), this translates to an average of 1.86 out of a maximum of 2 objects effectively tracked ($SD = 0.09$).¹ Data from the incorrect trials were not included in the subsequent ERP analyses.

Two separate N1 components were observed in this experiment. The posterior N1 occurred over lateral posterior electrode locations and peaked approximately 175 msec poststimulus for contralateral stimuli (Figure 1). The posterior N1 was measured as the average amplitude over a time window from 165 to 205 msec poststimulus. The anterior N1 occurred over lateral frontocentral electrode locations, peaked approximately 150 msec poststimulus, and was measured as the mean amplitude 120–160 msec poststimulus (Figure 2). Each analysis of the amplitude of these components was conducted on the average of four contiguous channels for each hemisphere. These channels showed the largest N1 amplitudes (averaged over target, distractor, and background conditions) for contralateral stimuli in the specified time windows. The channels used

for the posterior N1 were 58 (P9), 59 (P7), 65, and 66 for the left hemisphere and 92 (P8), 93 (P6), 97 (P10), and 98 (Tp8) for the right hemisphere. The channels used for the anterior N1 were 7, 13, 21 (Fc1), and 29 (Fc3) for the left hemisphere and 5, 113, 119 (Fc2), and 124 (F4) for the right hemisphere. Note that the P1 amplitudes observed in this experiment were very small and did not show any interpretable effect of probe location, so P1 analyses will not be included in the following results. All subsequent ANOVAs use the Greenhouse-Geisser correction when appropriate.

Posterior N1. Figure 1 shows that probes evoked a contralateral N1 over the posterior cortex that peaked at approximately 175 msec. The amplitude of the N1 appeared to be smallest for probes presented on distractor objects and comparable for target objects and the background. This was confirmed by a 2 (hemisphere: left vs. right) \times 3 (probe location: target, distractor, background) repeated measures ANOVA, which showed a significant effect of probe location [$F(2,34) = 5.83$, $MS_e = 0.342$, $p < .05$]. No other effects were reliable ($Fs < 1$, $ps > .38$). Follow-up *t* tests² indicated that N1 amplitudes were larger for target versus distractor probes [$t(17) = 3.45$, $p < .01$], as well as background versus distractor probes [$t(17) = 2.29$, $p < .05$]. Target and background probes did not differ [$t(17) = 0.60$, $p = .56$]. This pattern suggests that the N1 for distractor probes was suppressed, relative to target and background probes.

Experiment 1: Anterior N1

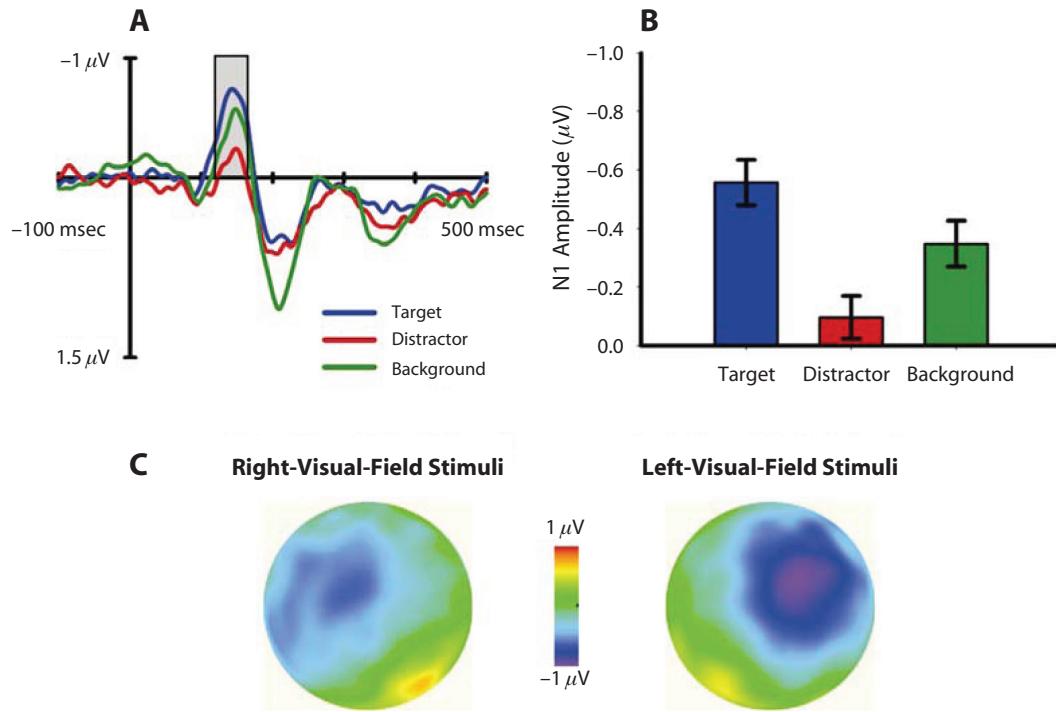


Figure 2. Anterior N1 results from Experiment 1. (A) Grand averaged ERP waveforms elicited by contralateral visual field stimuli. The shaded region corresponds to the analysis time window. (B) Mean N1 amplitude in the measurement time window. (C) Isopotential scalp maps for left- and right-visual-field stimuli 145 msec poststimulus. Note that target probes produce the largest amplitude N1, followed by background probes, and distractor probes produced the smallest N1 components. Error bars are ± 1 standard error.

Anterior N1. Anterior N1s are shown in Figure 2, where it appears that target probes elicited the largest N1, followed by background probes and then the distractor probes. Contralateral anterior N1 amplitude was evaluated with a 2 (hemisphere: left vs. right) \times 3 (probe location: target, distractor, background) repeated measures ANOVA, which showed a main effect of probe location [$F(2,34) = 13.95, MS_e = 0.157, p < .001$]. No other effects were reliable [hemisphere, $F(1,17) = 4.21, p = .06$; hemisphere \times probe location, $F(2,34) = 2.57, p = .11$]. Follow-up *t* tests showed a significant difference between target and distractor probes [$t(17) = 4.62, p < .001$] and between target and background probes [$t(17) = 2.33, p < .05$]. In addition, the distractor probes resulted in smaller N1s than did the background probes [$t(17) = 3.60, p < .01$]. If we consider the background probe to be a neutral baseline condition, the amplitude of the N1 increased for the target probes and decreased for the distractor probes. The results for the anterior N1 suggest that attention tracks objects during MOT by preferentially allocating visual attention to targets while simultaneously suppressing distractors.

Discussion

In this experiment, participants tracked two out of four objects while irrelevant probe flashes were presented on targets, distractors, or the background. The amplitude of

the N1 component of the ERP was used as a measure of the amount of visual attention allocated to each of these display areas. The posterior N1 showed a pattern consistent with the idea that distractors are suppressed during MOT. Target and background N1s were comparable in amplitude, and both were larger than distractor N1s. For the earlier occurring anterior N1, targets were larger than backgrounds, which, in turn, were larger than distractors. This pattern is consistent with simultaneous suppression of distractors and enhancement of targets.

The present results indicate that the suppression of distractors operated fairly early in processing, at least as early as the anterior N1, which had a peak latency of 150 msec. In addition, the anterior N1 showed that targets were enhanced relative to both distractors and the background, which is consistent with earlier reports that this component is larger for attended than for unattended flashes (Hillyard & Anllo-Vento, 1998). Apparently, the same process that is engaged when observers are cued to attend to a static display location is also used to track moving targets. This process appears to be one of enhancement of neural activity at early locations in the visual system for attended objects, together with suppression of unattended objects. This suppression pattern was also observed in the later posterior N1 at a latency of 175 msec.

These results are broadly consistent with Pylyshyn's (2006) finding that observers are more likely to detect

briefly presented probes occurring on targets and background areas, relative to distractor probes, a pattern interpreted as evidence that distractors are suppressed during MOT. According to Pylyshyn (2004), although indexes “attached” to targets are the main mechanism underlying object tracking, indexes may sometimes be lost to nearby distractors and are a significant source of tracking errors (see also O’Hearn et al., 2005). Suppression of distractors can then be seen as a useful *add-on* process that helps to ensure the success of the indexing mechanism when distractors are confusable and close to a target.

There are several reasons, however, to be cautious in accepting this pattern of results as supporting the claim that MOT generally involves suppression of distractors. First, the results of a similar study by Drew et al. (2009) suggest that tracked objects are enhanced during MOT without suppression of distractors. There are several differences between the Drew et al. study and the present one that might account for the different pattern of findings in the two studies. Our objects consisted of random dot textures moving on a moving random dot background, whereas Drew et al. used luminance-defined objects. Our probe flashes involved changing the luminance of the object itself, whereas their probe was a separate object occurring inside the contours of a hollow object. Finally, their displays included eight distractors (four moving and four stationary), as opposed to two distractors in our study. Additional research will be required to determine which of these differences in methodology might have been responsible for the different pattern of results.

Second, several experiments (Flombaum & Scholl, 2009; Flombaum et al., 2008; Pylyshyn, 2006; Pylyshyn et al., 2008; Experiments 2 and 3 in the present article) have shown that detection of background probes is often superior to *both* targets and distractors and this holds true even when observers are instructed to ignore the tracking task and just detect the probes (Pylyshyn, 2006). This raises the possibility that background probes may simply be more visible than probes occurring on objects and, therefore, may not constitute the ideal neutral condition for drawing inferences regarding suppression versus enhancement. Visibility differences might arise because of physical differences between object and background probes. For example, probes on objects will be surrounded by contours, which might produce more masking of the probe, as compared with those appearing on a blank background. In addition, background probes will generally be stationary, in contrast to the target and distractor objects, which move continuously. If moving objects are at a disadvantage, as compared with stationary ones (Flombaum & Scholl, 2009), then once again, background probes may have an advantage.

The present method of using moving texture objects on a moving texture background eliminates differences between object and background probes based on whether they are moving or stationary, as well as differences in masking by luminance contours. However, it certainly does not eliminate all potentially important differences between objects and background areas that might affect probe visibility, and so it is important to remain cautious

in drawing the inference that distractors are suppressed at early levels during MOT. We will continue to entertain the idea that MOT may rely on mechanisms of target enhancement, distractor suppression, or a combination of the two, while recognizing that additional research will be required to settle the issue. The most important inference to draw from these results is simply that early mechanisms of visual attention play a role in favoring tracked objects over distractors.

EXPERIMENT 2 Relevant Probes

The results of Experiment 1 support the claim that visual attention does play a role in MOT. However, it is not clear whether the allocation of visual attention is *necessary* for MOT or is simply an optional strategy. Experiment 2 addressed this issue by providing observers with an incentive to spread their attention over the entire display during tracking. In this experiment, the probe flashes were task relevant, since the observers were asked to detect occasional target probes embedded in the stream of probe flashes occurring on tracked objects (target objects), distractor objects, and the background. If allocation of visual attention to target objects is obligatory for accurate object tracking, we should see superior detection of target probes when they occur on tracked objects, relative to distractors and backgrounds. In addition, we should see a pattern of N1 amplitudes similar to that observed in Experiment 1. On the other hand, if the allocation of visual attention observed in Experiment 1 was optional, observers would have no particular incentive to allocate visual attention to tracked objects (targets), and we should observe N1s of comparable amplitude for probes occurring anywhere in the display.

The addition of relevant target probes also allowed us to measure the P300 component, which is a later, post-perceptual component that has been associated with consolidation of information into memory (Donchin, 1981; Donchin & Coles, 1988) and is affected by attention (for reviews, see Coull, 1998; Polich & Kok, 1995). The P300 is also sensitive to the probability of various categories of eliciting stimuli, such that low-probability events produce larger amplitude P300 components than do high-probability events (Duncan-Johnson & Donchin, 1977; Kutas, McCarthy, & Donchin, 1977). If probes occurred equally often on targets, distractors, and the background, as in Experiment 1, and observers categorized the probe events as occurring on either objects or the background, background probes would occur only one third of the time and could produce a large P300 simply because they are rarer than the other events. We tested this by running two versions of the experiment. Experiment 2A was similar to the procedure used in Experiment 1, in which probes occurred equally often on targets, distractors, and the background. In Experiment 2B, half of the probes occurred on the background, and the other half were divided equally among targets and distractors. To foreshadow our results, no probability effects were observed on the P300, so we were able to combine the data from both versions of the

experiment for added power in detecting any potentially small N1 effects.

Method

Participants. Forty-one right-handed, neurologically normal volunteers (ages, 18–32 years) participated in this experiment (Experiment 2A, $n = 27$; Experiment 2B, $n = 14$). With the exception of author M.M.D. (Experiment 2A), all were naive as to the purpose of the experiment and were paid \$10/h for their participation. All the participants reported normal or corrected-to-normal acuity and provided informed consent. Four participants from Experiment 2A were excluded from the analyses for the following reasons. One participant exhibited exceptionally poor probe detection performance (fewer than 20% of the probes were detected). One participant never detected background or distractor probes, which we took as evidence that the participant did not follow the instruction to respond to probes regardless of where they occurred. One participant showed excessive alpha in the ERP recording. One participant was excluded because too many channels were bad in the ERP recording (more than 10 channels were marked bad in all of the segments).

Stimuli and Procedure. These experiments were identical to Experiment 1, except as noted below. In Experiment 1, all the probe flashes were irrelevant to the observer, but in these experiments, 25% of the probe flashes required a response (*response probes*). Response probes were identical to the irrelevant probes, except that there was a small gray square (0.11°; 50.09 cd/m²) in their centers. In Experiment 2A, these response probes were evenly distributed over targets, distractors, and the background. In Experiment 2B, half of the probes occurred on the background, and the other half of the probes were equally divided among targets and distractors. In both cases, each probe type occurred equally often in both visual fields. All the probes were presented in a random order, but the presentation was constrained so that the first probe on a trial was an irrelevant probe and response probes were separated by at least 1 irrelevant probe. A total of 12 (Experiment 2A) or 16 (Experiment 2B) probes were presented during each MOT trial. The number of response probes presented during each MOT trial varied from 2 to 4 (Experiment 2A) or 3 to 5 (Experiment 2B). Responses that occurred 200–900 msec after the presentation of a response probe were counted as hits. If no response was made within this time window, a miss was recorded. Any other responses were counted as false alarms.

Each experiment consisted of 108 trials that lasted 10 sec (Experiment 2A) or 13 sec (Experiment 2B). As a result, the participants viewed a total of 162 irrelevant probes and 54 response probes per condition, created by crossing visual field (left vs. right) and probe location (target, distractor, or background) in Experiment 2A. In Experiment 2B, the participants viewed 162 irrelevant and 54 response probes in each visual field for target and distractor probe locations,

and 324 irrelevant and 108 response probes occurred on the background in each visual field. The observers responded to the *target* probes by clicking a mouse button regardless of whether the probe appeared on a tracked object, a distractor, or the background. The participants were instructed to prioritize MOT accuracy over probe detection accuracy. In addition to feedback about tracking accuracy and eye movements/blinks, the participants received feedback about the number of hits, misses, and false alarms that were recorded at the end of each trial.

Results

Tracking accuracy was high in both versions of the experiment. In Experiment 2A, the participants correctly selected both target objects on 89.6% of the trials ($SD = 5.7$), which translates to an average of 1.80 ($SD = 0.10$) objects effectively tracked using the formula from Pylyshyn and Annan (2006). In Experiment 2B, the participants correctly selected both target objects on 87.6% of the trials ($SD = 4.9$), which translates to an average of 1.76 ($SD = 0.09$) objects effectively tracked. It is worth noting that the addition of the probe detection task in these experiments produced a small but reliable decrease in accuracy on the MOT task, relative to Experiment 1 [$t(53) = 2.49, p < .05$]. Data from the incorrect tracking trials were not included in any analyses of probe detection or ERPs.

Probe detection. On average, the participants correctly detected 70.1% of the response probes ($SD = 13.9$) in Experiment 2A and 59.5% ($SD = 19.9$) in Experiment 2B. False alarm rates were very low and averaged only 0.21 false alarms per MOT trial ($SD = 0.13$) in Experiment 2A and 0.19 false alarms per MOT trial ($SD = 0.18$) in Experiment 2B. As a result, false alarms were not included in the probe detection analysis. Probe hit rates are shown in Figure 3, where it appears that for both versions of Experiment 2, performance was highest for background probes, intermediate for targets, and lowest for distractors. This was confirmed by a 2 (experiment: 2A vs. 2B) \times 2 (visual field) \times 3 (probe location: target, distractor, background) ANOVA. A main effect of experiment indicated that probes were detected more frequently in Experiment 2A than in Experiment 2B [$F(1,35) = 4.87, MS_e = 0.139, p < .05$]. Importantly, though, this factor did not interact with any other factors ($F_s < 2.92, ps > .09$). A significant

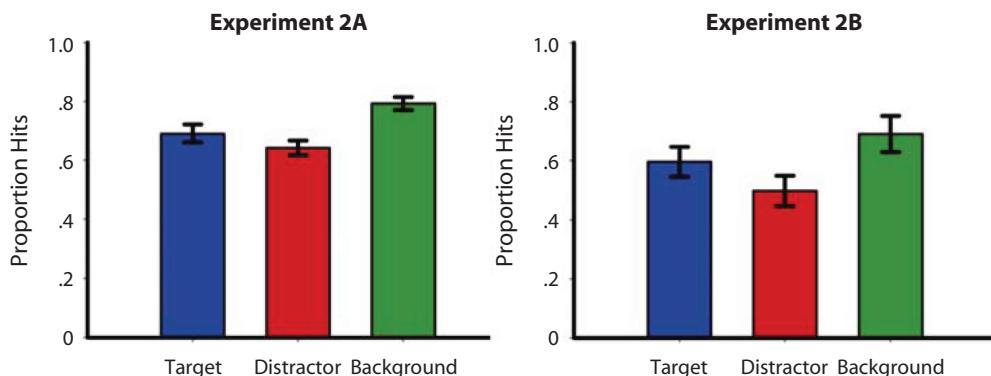


Figure 3. Probe hit rates in Experiment 2. Background probes were detected more readily than target probes, and distractor probes were detected least often. Note that this pattern holds for both versions of Experiment 2. Error bars represent ± 1 standard error.

effect of visual field indicated that probes were detected more often in the right visual field ($M = 69.3\%$, $SD = 16.4$) than in the left visual field ($M = 63.8\%$, $SD = 17.1$) [$F(1,35) = 9.08$, $MS_e = 0.014$, $p < .01$]. More interestingly, a main effect of probe location [$F(2,70) = 72.31$, $MS_e = 0.010$, $p < .001$] indicated that performance was highest for background probes [background vs. target, $t(36) = 5.73$, $p < .001$; background vs. distractor, $t(36) = 12.10$, $p < .001$], followed by target probes [target vs. distractor, $t(36) = 6.56$, $p < .001$]; distractor probes were detected least often. This pattern is consistent with the existing behavioral work, which has been interpreted as evidence for distractor suppression in MOT (e.g., Pylyshyn, 2006; Pylyshyn et al., 2008), although, as we pointed out earlier, the superiority of background probes relative to target probes is not predicted by distractor suppression.

P300. The P300 was observed over central-parietal electrode locations and peaked approximately 475 msec after onset of the probe (Figure 4). The mean amplitude was measured at Channels 54 (P1), 55 (Cpz), 62, and 80 (P2), using a 425- to 525-msec window following the onset of the probe. Note that only the data from the response probes were included in the following analysis. A 2 (experiment: 2A vs. 2B) \times 2 (visual field: left vs. right) \times 3 (probe location: target, distractor, background) repeated measures ANOVA did not show any reliable effects or interactions with the experiment factor ($F_s < 3.74$, $ps > .06$). There was a main effect of visual field [$F(1,35) = 4.78$, $MS_e = 1.21$, $p < .05$], indicating that right-visual-field stimuli produced larger P300s than did left-visual-field stimuli (right visual field, $M = 3.29 \mu V$, $SD = 1.94$; left visual field, $M = 2.88 \mu V$, $SD = 2.18$).

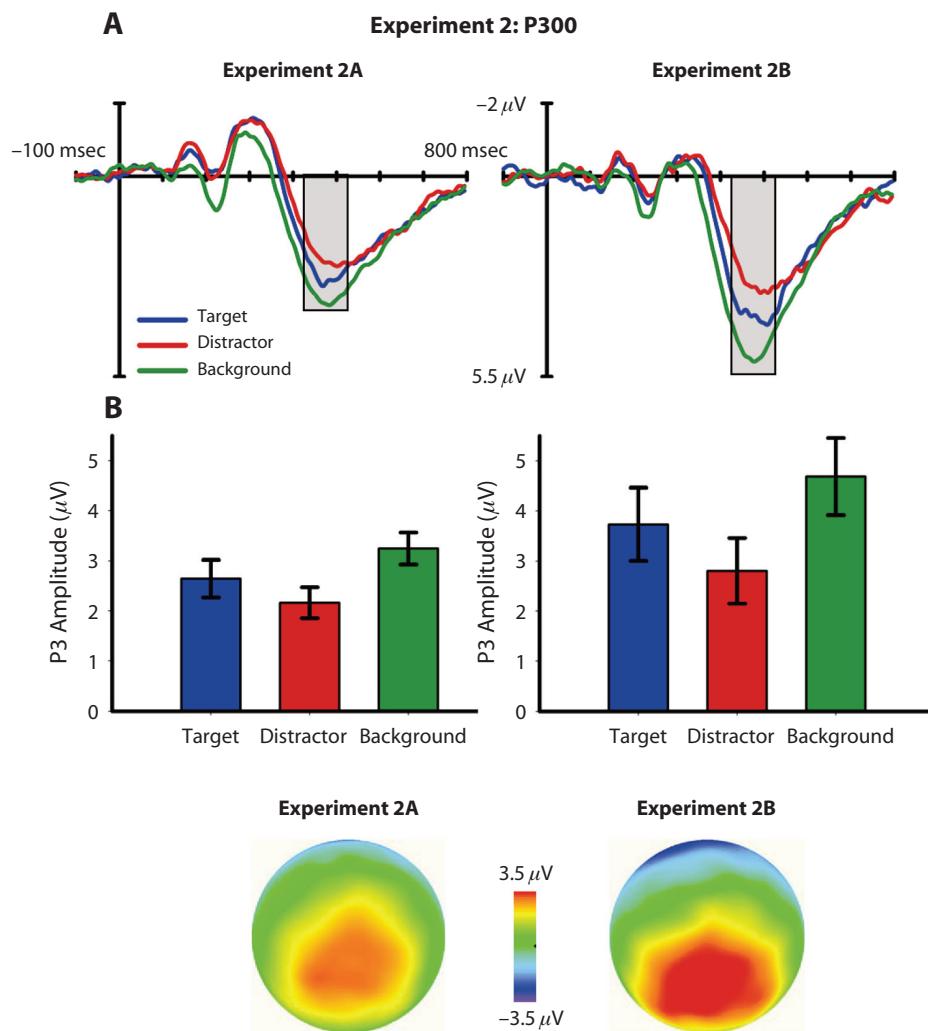


Figure 4. P300 results from Experiment 2. (A) Grand averaged ERP waveforms for each version of Experiment 2. The shaded region corresponds to the analysis time window. (B) Mean P300 amplitude in the measurement time window. (C) Isopotential scalp maps for target probes in each version of Experiment 2 at 475 msec poststimulus. Background probes produced the largest amplitude P300, followed by target probes, and distractor probes produced the smallest P300s. Note that this pattern holds for both versions of Experiment 2 and mirrors the behavioral probe detection rates. Error bars are ± 1 standard error.

Experiment 2: Posterior N1

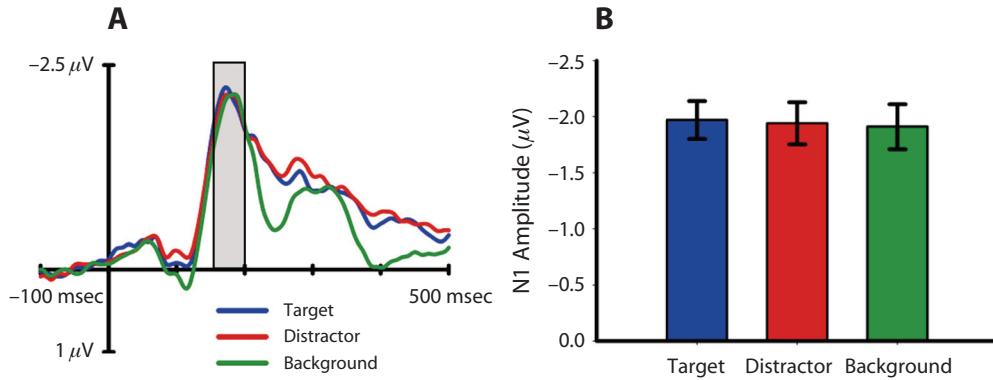


Figure 5. Posterior N1 results from Experiment 2. (A) Grand averaged ERP waveforms elicited by contralateral visual field stimuli. The shaded region corresponds to the analysis time window. (B) Mean N1 amplitude in the measurement time window. Note that probe location (i.e., target, distractor, or background) does not seem to affect N1 amplitude. Error bars are ± 1 standard error.

There was also a main effect of probe location [$F(2,70) = 30.62, MS_e = 1.261, p < .001$]; no interactions involving probe location were reliable ($F_s < 2.24, ps > .11$). Background probes produced the largest P300, followed by targets, with distractors producing the smallest P300 [target vs. distractor, $t(36) = 3.53, p < .01$; target vs. background, $t(36) = 3.86, p < .001$; distractor vs. background, $t(36) = 7.45, p < .001$]. P300 amplitude, shown in Figure 4, nicely mirrors the behavioral probe detection results reported above and elsewhere (Pylyshyn, 2006). In addition, the comparability of results across the two versions of Experiment 2 appears to rule out probability confounds as an explanation for the larger P300s associated with background probes. We conclude that at the level of the P300, targets and distractors are differentiated and that this pattern is consistent with behavioral data interpreted as evidence for distractor suppression in MOT (e.g., Pylyshyn, 2006; Pylyshyn et al., 2008), although, again, the larger amplitude P300 for background probes is not predicted by distractor suppression.

Posterior N1. In order to maximize our statistical power, the following N1 analyses treat the data from the two versions of Experiment 2 as if they came from a single experiment with 37 participants. Figure 5 shows amplitude plots of the N1 as a function of probe condition. The N1 appears as a contralateral negativity over lateral posterior electrode sites with a latency of approximately 180 msec poststimulus. Its amplitude is remarkably similar across the different probe locations. N1 amplitude was quantified as the average voltage over a time window from 160 to 200 msec poststimulus for Contralateral Electrodes 59 (P7), 60, 65, and 66 for the left hemisphere and 85, 86, 92 (P8), and 93 (P6) for the right hemisphere. Note that only the data from the irrelevant probes were used in the following analysis (i.e., the response probes were excluded). As in Experiment 1, N1 amplitude was evaluated with a 2 (hemisphere: left vs. right) \times 3 (probe location: target, distractor, background) repeated measures ANOVA. No reliable main effects or interactions were observed

($F_s < 1, ps > .41$). In other words, it appears that, contrary to the results of Experiment 1, there was no reliable effect of probe location on posterior N1 amplitude.

Anterior N1. The anterior N1 was observed over lateral frontocentral electrode locations and peaked approximately 155 msec poststimulus (Figure 6). The anterior N1 was not as clearly differentiated in terms of scalp topography as in Experiment 1, partly because it overlapped with the posterior N1, which had a slightly earlier onset in this experiment. Consequently, we used the same electrodes as in Experiment 1 for analysis of this component. The anterior N1 was defined as the mean amplitude in a 135- to 175-msec poststimulus window measured at electrodes contralateral to the probe flash. Note that only the data from the irrelevant probes were used in the following analysis (i.e., the response probes were excluded). A 2 (hemisphere: left vs. right) \times 3 (probe location: target, distractor, background) repeated measures ANOVA did not show any reliable main effects or interactions ($F_s < 2.13, ps > .13$).

Power. The failure to find any effect of probe location on either anterior or posterior N1 amplitude stands in contrast to the results of Experiment 1, which showed robust effects of probe location on both N1 components. It seems unlikely that the lack of significant effects can simply be attributed to a lack of power. Using G*Power 3.0.10 (Faul, Erdfelder, Lang, & Buchner, 2007), we determined that there was sufficient power (power = .80, $\alpha = .05$) to detect main effects of probe location with effect sizes of .0353 and .0158 for the anterior and posterior N1 components, respectively (effect sizes are partial η^2 values). The effect sizes in Experiment 1 were as follows: partial $\eta^2 = .45$ for the anterior N1 and partial $\eta^2 = .24$ for the posterior N1. So, we had reasonable power to detect effects less than one twelfth the size of those observed in the previous experiment.

Discussion

Our behavioral results provide a replication of those in previous experiments (Flombaum et al., 2008; Pylyshyn,

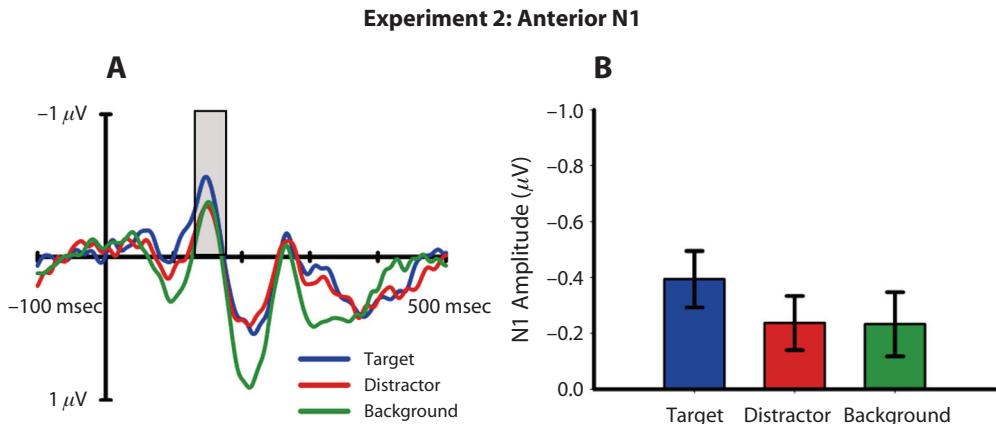


Figure 6. Anterior N1 results from Experiment 2. (A) Grand averaged ERP waveforms elicited by contralateral visual field stimuli. The shaded region corresponds to the analysis time window. (B) Mean N1 amplitude in the measurement time window. Note that probe location (i.e., target, distractor, or background) does not seem to affect N1 amplitude. Error bars are ± 1 standard error.

2006; Pylyshyn et al., 2008), showing that the ability to detect target probes during MOT depends on probe location: Accuracy was highest with background probes, intermediate for targets, and lowest for distractors. This pattern has been taken as evidence for distractor suppression (Pylyshyn, 2006; Pylyshyn et al., 2008), although the finding that background probes are detected more frequently than target probes appears to require an explanation over and above the idea of distractor suppression. The amplitude of the P300 component corresponded closely to the behavioral detection data, which perhaps is not surprising in the light of previous work showing that P300 amplitude is closely related to hit rate in signal detection paradigms (Imai & Tsuji, 2004).

The effects of probe location on N1 amplitude, however, were quite different. Contrary to the results of Experiment 1, we found no effects of probe location on the amplitudes of either the anterior or the posterior N1 component. In other words, the addition of task-relevant probes in Experiment 2 seems to have altered the distribution of visual attention reflected in the N1 components during MOT. We assume that in an attempt to detect target probes occurring anywhere in the display, the observers distributed their visual attention over the entire display, rather than restricting it to areas containing tracked objects. A distributed mode of allocating spatial attention would account for the lack of N1 differences associated with different probe locations.

Another possible explanation of the lack of N1 effects in this experiment is that attention was switched periodically between the tracking and probe detection tasks throughout the trial. On the surface, this appears plausible, since it has been shown recently that MOT can be periodically interrupted without degrading tracking performance. For example, Horowitz, Birnkrant, Fencsik, Tran, and Wolfe (2006) demonstrated good tracking performance even when the tracking display was hidden for periods of up to 0.5 sec, during which the objects continued to move.

Could it be that in the present experiment, visual attention was involved in the tracking task but periodically switched to a more broadly tuned state in order to perform the probe detection task? This sort of account seems unlikely, because it still predicts that probes appearing on targets during periods when attention was allocated to the tracking task should have resulted in larger N1s for targets than for distractors, and that is not what we found. Depending on what fraction of time attention is allocated to targets, the N1 difference would be smaller than the one observed in Experiment 1, but the power analysis shows that there was sufficient power to detect effects more than 12 times smaller than those observed in Experiment 1, in which there was no reason to switch attention from the tracked objects.

Another possibility is that the observers did not need to periodically switch to a different attentional state in anticipation of a probe flash, because flashes in Experiment 2 were able to capture attention automatically. In other words, the observers attended to target objects in the MOT task while also adopting a display-wide attentional set for sudden onsets so that their attention could be rapidly allocated to probes (Most, Scholl, Clifford, & Simons, 2005). If this attention switch occurred rapidly enough to enhance the N1 component, essentially all probes, regardless of their locations, would be “attended” and produce equivalent N1 components.

This explanation seems unlikely because it appears that the N1 reflects the state of attention existing at the time of probe onset and not changes in that allocation that occur in response to the probe onset. Attention switching in response to a sudden onset is reflected in a different component (the N2PC; Luck & Hillyard, 1994b), which has a latency just beyond that of the N1. In addition, others have reported N1 attention effects in paradigms that might have encouraged the same sort of attentional set by using relevant target events that could occur in an attended location or elsewhere (Eimer, 1997; Mangun & Hillyard, 1988, 1991).³

For example, Mangun and Hillyard (1988) told participants that their primary task was to respond to infrequent target events that occurred in a prespecified attended location. As a secondary task, the participants were instructed to also respond to target events that occurred in either of two “unattended” locations. In this case, as in the present experiment, it would have been advantageous for the participants to adopt an attentional set whereby the onset stimuli could capture attention in order to discriminate targets from nontargets at all locations in the display. However, the results still showed larger P1/N1 components for the stimuli in attended locations, relative to the unattended locations. So, at least for the time being, it seems likely that the lack of N1 effects in this experiment is attributable to a broad distribution of attention, rather than to an attention-switching strategy. Nonetheless, we acknowledge that the ability of observers to perform the MOT task with asynchronous access to the display makes attention switching a potentially viable strategy in dealing with dual-task situations and will require further research before it can definitively be ruled out as an explanation for the striking ability of MOT to coexist with other attention-demanding activities.

In sum, the results of this experiment suggest that suppression is a flexible process that can be applied at different processing levels, depending on the task (Vogel et al., 2005). In Experiment 1, in which probe flashes were irrelevant, it made sense to filter out distractors as soon as possible, and differences between target and distractor probes were observed in both anterior and posterior N1 components. However, in Experiment 2, in which probes were relevant and occasionally required detection responses, early suppression would mean a high miss rate for distractor probes. In this case, it makes sense to allow the probes access to higher stages of analysis that might be able to make a preliminary determination of whether the probe is likely to be a target that requires a response. This finding also suggests that visual attention is not a *necessary* process for MOT (at least when tracking load is fairly light), because tracking can apparently occur without differential visual attention’s being allocated to targets and distractors. Furthermore, these results suggest that caution is warranted in interpreting behavioral probe detection rates during MOT in terms of the allocation of visual attention (especially when tracking load is light), since selectivity in probe detection rates was observed in the absence of any N1 differences.

EXPERIMENT 3 Difficult Tracking and Relevant Probes

The results of Experiment 1 showed that visual attention was engaged during the MOT task, and Experiment 2 showed that accurate tracking could still occur when attention was distributed over the entire display. These results suggest that the investment of visual attention in tracking that we observed in the first experiment was *optional* and not required for accurate tracking. However, it is important to consider the generality of this finding. It may be that visual attention is not important for *easy*

tracking but plays an important role when tracking is *difficult*. In Experiment 3, we increased tracking difficulty by requiring observers to track three targets among three distractors. As in Experiment 2, probes were task relevant. If visual attention is needed for tracking under more difficult circumstances, we might be able to reinstate the effects of attention on the N1 components that we observed in Experiment 1.

Method

Participants. Nineteen right-handed, neurologically normal volunteers (ages, 18–32 years) participated in this experiment. All were naive as to the purpose of the experiment and were paid \$10/h for their participation. All the participants reported normal or corrected-to-normal acuity and provided informed consent. Three participants were excluded from the analyses for the following reasons. One participant exhibited exceptionally poor probe detection performance (fewer than 1% of the probes were detected). Two participants showed excessive alpha in the ERP recording.

Stimuli and Procedure. Since the results of Experiment 2 did not show any effects of the probability manipulation, we did not include any such manipulation in this experiment. This experiment was identical to Experiment 2A, except as noted below. In Experiment 2, participants tracked two out of four objects. In order to make the tracking task more difficult, the participants in the present experiment tracked three out of six objects. We expected that this increase in difficulty would lead to more tracking errors, so this experiment consisted of 120 MOT trials (rather than 108, as in Experiment 2) in order to compensate for the loss of probe presentations due to tracking errors. As a result, each participant viewed a total of 180 irrelevant and 60 response probes per condition, created by crossing visual field (left vs. right) with probe location (target, distractor, or background).

Results

Tracking accuracy in this experiment was still fairly good but was lower than that in the previous experiment [$t(51) = 5.14, p < .001$]. In this experiment, the participants correctly selected both target objects on 79.5% of the trials ($SD = 7.5$), as compared with 88.9% of the trials ($SD = 5.4$) in Experiment 2. This drop-off in performance is sensible given that, in this experiment, the participants were required to track an additional object in the presence of an additional distractor. Using the formula provided by Pylyshyn and Annan (2006; see also Hulleman, 2005, for a similar formula), the observers effectively tracked an average of 2.66 out of a maximum of three objects in this experiment ($SD = 0.13$).

Probe detection. On average, the participants correctly detected 52.3% of the response probes ($SD = 19.1$), as compared with 66.6% ($SD = 16.0$) in Experiment 2, indicating that the increased difficulty of the tracking task resulted in poorer probe detection performance [$t(51) = 2.81, p < .01$]. As in the previous experiment, false alarm rates were very low and averaged only .19 false alarms per MOT trial ($SD = .16$). As a result, false alarms were not included in the probe detection analysis.

Probe hit rates are shown in Figure 7, where it appears that performance was highest for background probes, intermediate for targets, and lowest for distractors. This was confirmed by a 2 (visual field) \times 3 (probe location: target, distractor, background) ANOVA. A main effect of probe location [$F(2,30) = 26.39, MS_e = 0.011, p < .001$]

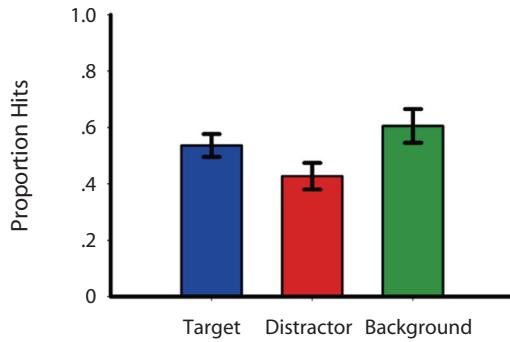


Figure 7. Probe hit rates in Experiment 3. Although probe hit rates are lower than in Experiment 2, the same pattern was observed: Background probes were detected more readily than target probes, and distractor probes were detected least often. Error bars represent ± 1 standard error.

indicated that performance was highest for background probes [background vs. target, $t(15) = 2.44, p < .05$; background vs. distractor, $t(15) = 7.79, p < .001$], followed by target probes [target vs. distractor, $t(15) = 4.85, p < .001$]; distractor probes were detected least often. No other effects or interactions were reliable ($Fs < 1.2, ps > .33$). This pattern is consistent with the existing behavioral work (e.g., Pylyshyn, 2006; Pylyshyn et al., 2008; the present Experiment 2), which has been interpreted as evidence for distractor suppression in MOT.

P300. The P300 was observed over central-parietal electrode locations and peaked approximately 500 msec poststimulus (Figure 8). The mean amplitude was measured using a 450- to 550-msec window following the onset of the probe at Channels 54 (P1), 61 (Po3), 62, and 79 (Po4). Note that only data from the response probes were included in the following analysis. A 2 (visual field: left vs. right) \times 3 (probe location: target, distractor, background) repeated measures ANOVA showed a main effect of visual field [$F(1,15) = 9.92, MS_e = 0.738, p < .01$],

indicating that right-visual-field stimuli produced larger P300s than did left-visual-field stimuli (right visual field, $M = 2.99 \mu V, SD = 2.05$; left visual field, $M = 2.43 \mu V, SD = 1.91$). There was also a main effect of probe location [$F(2,30) = 5.30, MS_e = 1.833, p < .05$], but the interaction was not reliable ($F < 1, p > .56$). Background and target probes produced similar amplitude P300 components [$t(15) = 1.33, p = .20$] that were both larger than the distractor P300 [target vs. distractor, $t(15) = 2.42, p < .05$; distractor vs. background, $t(15) = 3.25, p < .01$; see Figure 8]. As in Experiment 2, probe location affects the amplitude of the P300, and the pattern is consistent with distractor suppression.

Posterior N1. Figure 9 shows amplitude plots of the N1 as a function of probe condition. The N1 appears as a contralateral negativity over lateral posterior electrode sites with a latency of approximately 180 msec poststimulus. N1 amplitude was quantified as the average voltage over a time window from 160 to 200 msec poststimulus for Contralateral Electrodes 52 (P5), 59 (P7), 60, and 66 for the left hemisphere and 78, 86, 92 (P8), and 93 (P6) for the right hemisphere. Note that only the data from the irrelevant probes were used in the following analysis (i.e., the response probes were excluded). N1 amplitude was evaluated with a 2 (hemisphere: left vs. right) \times 3 (probe location: target, distractor, background) repeated measures ANOVA. A main effect of hemisphere indicated that larger amplitude N1 components were observed at left-hemisphere than at right-hemisphere electrode locations [$F(1,15) = 4.85, MS_e = 2.674, p < .05$]. More important, a main effect of probe location was observed [$F(2,30) = 4.65, MS_e = 0.485, p < .05$]. As can be seen in Figure 9, target and background probes produced comparable N1 components [$t(15) = 1.15, p = .27$], and target probes produced larger N1s than did distractor probes [$t(15) = 4.68, p < .001$]—especially in the left hemisphere, although the interaction was not reliable [$F(2,30) < 1, p = .39$]. This pattern of data is partially consistent with the suppression of distractors at the level of the N1, except

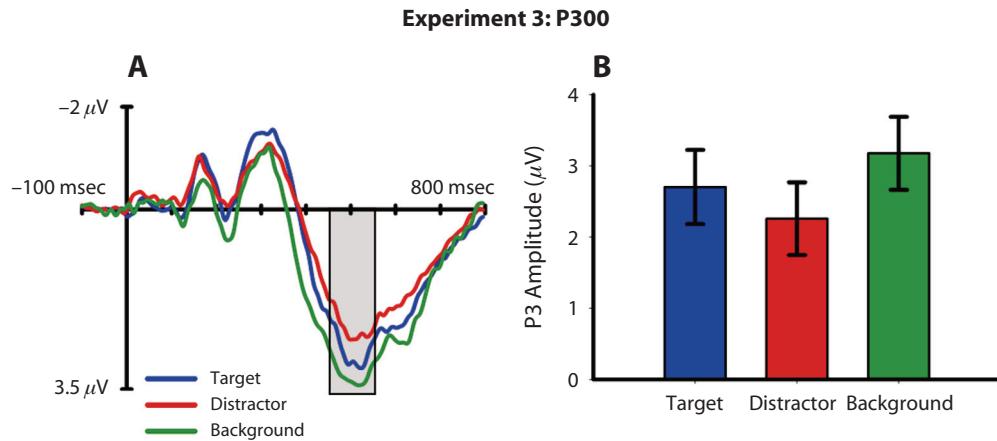


Figure 8. P300 results from Experiment 3. (A) Grand averaged ERP waveforms for Experiment 3. The shaded region corresponds to the analysis time window. (B) Mean P300 amplitude in the measurement time window. Background and target probes produced similar amplitude P300s that were larger than those produced by distractor probes. Error bars are ± 1 standard error.

Experiment 3: Posterior N1

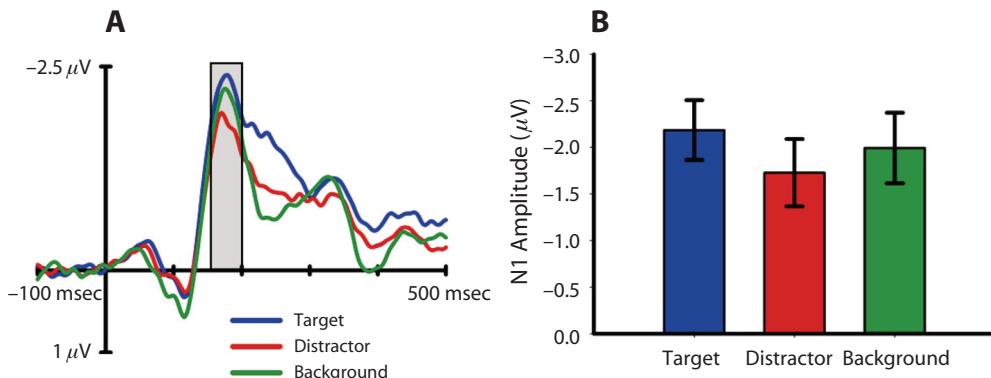


Figure 9. Posterior N1 results from Experiment 3. (A) Grand averaged ERP waveforms elicited by contralateral visual field stimuli. The shaded region corresponds to the analysis time window. (B) Mean N1 amplitude in the measurement time window. Target probes elicited larger amplitude N1 components than did distractor probes. Error bars are ± 1 standard error.

that N1 amplitudes for background and distractor probes were not reliably different [$t(15) = 1.52, p = .15$]. However, it is clear that, as in Experiment 1 and contrary to the results of Experiment 2, posterior N1 amplitude is larger for targets than for distractors, suggesting the involvement of visual attention in MOT.

Anterior N1. The anterior N1 was observed over lateral frontocentral electrode locations and peaked approximately 150 msec poststimulus (Figure 10). As in Experiment 2, the anterior N1 was not as clearly differentiated in terms of scalp topography as in Experiment 1, partly because of overlap with the posterior N1, which began slightly earlier in this experiment than in Experiment 1. Consequently, we used the same electrodes as in Experiment 1 for analysis of this component. The anterior N1 was defined as the mean amplitude in a 130- to 170-msec poststimulus window measured at electrodes contralateral to the probe flash. Note that only the data from the irrelevant probes were used in the following analysis (i.e., the response probes were excluded). A 2 (hemisphere: left vs. right) \times 3 (probe location: target, distractor, background) repeated measures ANOVA did not show a significant effect of hemisphere [$F(1,15) = 3.92, MS_e = 0.286, p = .07$] or an interaction [$F(2,30) = 0.267, MS_e = 0.191, p = .70$]. More interestingly, a significant effect of probe location [$F(2,30) = 11.63, MS_e = 0.205, p < .001$] indicated that target probes produced the largest N1s [target vs. distractor, $t(15) = 4.49, p < .001$; target vs. background, $t(15) = 4.61, p < .001$], whereas background and distractor N1 amplitudes did not differ [$t(15) = 0.33, p = .74$]. In other words, the target N1 was increased relative to both distractors and the background. Thus, contrary to the results in Experiment 2 but consistent with those in Experiment 1, visual attention seems to be involved in the tracking task. In this experiment, the effect of probe location on anterior N1 amplitude is one of pure target enhancement, which is consistent with the findings of Drew et al. (2009).

P2 component. The observant reader may have noticed another component in our anterior recordings from

all three experiments (Figures 2, 6, and 10) that we have ignored until now. Following the anterior N1, there is a positive component (the P2) that is larger for background probes than for either target or distractor probes, which are similar to each other. We quantified this as the average amplitude of Electrodes 6 (Fcz), 7, 107, and 129 (Cz) in the time window of 195–235 msec. The average amplitude across the three experiments was 0.62 μ V for targets, 0.63 μ V for distractors, and 1.11 μ V for background probes. A repeated measures ANOVA revealed a significant main effect of probe location [$F(2,136) = 27.65, MS_e = 0.36, p < .001$], and no other effects or interactions were reliable ($Fs < 1.8, ps > .14$). Follow-up *t* tests revealed significant differences between the background probes and both kinds of object probe: target probes [$t(70) = 6.88, p < .001$] and distractor probes [$t(70) = 6.54, p < .001$]. There was no significant difference between targets and distractors [$t(70) = 0.07, p = .94$]. In addition, the effect of probe location was similar across the three experiments, since the experiment \times probe location interaction was not significant.

These results show that background probes are indeed different from probes that occur on objects and that this difference appears in a time interval that is later than the N1. We will examine the implications of these P2 differences in the General Discussion section.

Discussion

In this experiment, we made the tracking task more difficult by adding a target and a distractor to the display. The increase in tracking difficulty was accompanied by a return of attention effects on both the anterior and posterior N1 components. The anterior N1 showed a pattern of pure *target enhancement* in which N1 amplitude to target probes was greater than that to both distractors and background probes, which did not differ. This can be contrasted to the anterior N1 effects in Experiment 1, which showed a combination of target enhancement and distractor suppression, and Experiment 2, which showed no probe loca-

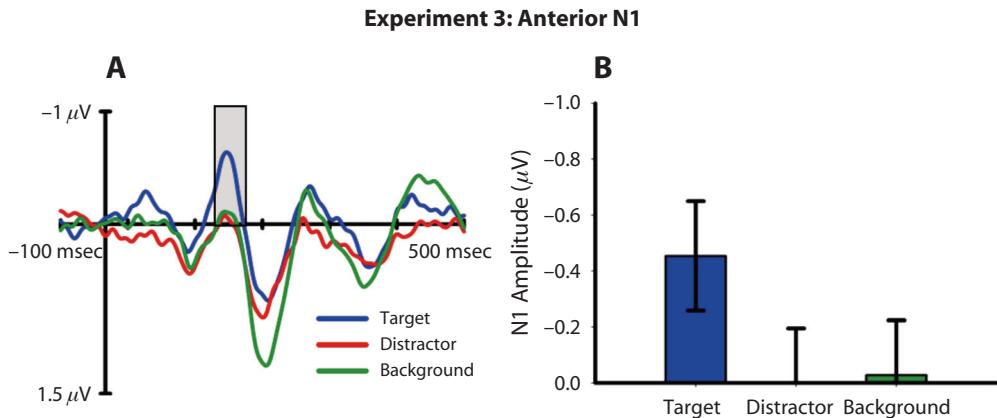


Figure 10. Anterior N1 results from Experiment 3. (A) Grand averaged ERP waveforms elicited by contralateral visual field stimuli. The shaded region corresponds to the analysis time window. (B) Mean N1 amplitude in the measurement time window. Background and distractor probes elicited similar amplitude N1 components that were smaller than those elicited by target probes. Error bars are ± 1 standard error.

tion effects at all. For the posterior N1 component, target and background probes were equivalent and larger than distractor N1s, although the background probe was not significantly different from the distractor. Thus, we have weak evidence that there may be distractor suppression at the level of the posterior N1, as we found in Experiment 1. In any case, the important point is that both the anterior and posterior N1 components once again showed attention effects. As was the case in Experiment 2, the P300 component showed a pattern that mirrored the behavioral detection results, with P300s being larger for both background and target probes than for distractors.

This pattern of results suggests that when observers are required to detect probes occurring throughout the display and displays are relatively sparse, as they were in Experiment 2, they adopt an attention allocation policy in which attention is spread broadly over the display, perhaps because early suppression might result in unacceptably low accuracy in detecting probes. When the difficulty of the primary tracking task is increased, observers respond by increasing the allocation of *early* visual attention processes to the tracked targets, resulting in larger N1s to target probes than to distractors. In a sense, the pattern observed in Experiment 2 represents a compromise between maintaining good performance on a difficult object-tracking task while continuing to detect target probes throughout the display at a reasonable level. Clearly, across these three experiments, we are looking at different points on the trade-off function defining joint performance on the two tasks (Alvarez, Horowitz, Arsenio, DiMase, & Wolfe, 2005), and an important task for future research will be to assess the full attention-operating curve, using both early and late ERP components.

GENERAL DISCUSSION

Summary of “Early” N1 Attention Effects

Our goal was to use the amplitude of the N1 component of the ERP to probes occurring in different parts of a display to assess the role of visual attention in MOT. In

order to simplify our summary of results, we will ignore for the moment the issue of whether N1 amplitude effects reflect distractor suppression or target enhancement and will consider a difference between target and distractors as evidence of an *attention effect*. The issue of distractor suppression will be considered in detail below.

In Experiment 1, the observers tracked two objects among two distractors, and we found that irrelevant probes on tracked objects produced larger anterior and posterior N1s than did probes on distractors, which we take as evidence that the observers allocated more visual attention to tracked objects than to distractors. This finding is similar to that in Drew et al. (2009), in which luminance-defined objects and denser displays (two targets and eight distractors) were used.

In Experiment 2, observers performed the same tracking task while trying to detect occasional targets in the stream of probes. Although tracking performance was similar to that observed in Experiment 1, making the probes relevant had a dramatic effect on the N1 component, with equivalent amplitudes observed for target, distractor, and background probes. We interpret this as evidence that observers distributed visual attention over the entire display in an effort to detect probes on distractors and background areas, as well as on targets. The fact that tracking accuracy remained high suggests that the differential allocation of visual attention to targets versus distractors observed in Experiment 1 was not *necessary* for tracking.

The results of Experiment 2 suggest that visual attention is not a *necessary* component of MOT. However, it might play an important role with higher tracking loads, as suggested by theories such as Lavie’s load theory (Lavie, 2005; Lavie, Hirst, de Fockert, & Viding, 2004), which holds that selectivity tends to increase with increasing perceptual load. To test this, in Experiment 3, we required observers to track three targets in a display of six objects. Note that this also increased display density so that, on average, targets passed closer to distractors, and this should increase the perceptual load of the tracking task. As in Ex-

periment 2, the probes were relevant. In this case, robust attention effects were once again observed with the larger N1 amplitude associated with target probes, as compared with distractors. This apparent shift of visual attention to the tracking task was accompanied by a drop in accuracy on the probe task, as compared with Experiment 2, which we view as evidence for a competition for visual attention between the two tasks, a competition that was not apparent for the sparser displays used in Experiment 2.

Visual Attention Versus Indexing

These results have important implications for understanding the role of visual attention in MOT. Experiment 2 showed that observers can sometimes track multiple targets without differentially allocating *visual attention* to targets versus distractors. Of course, this conclusion needs to be qualified by the usual concerns regarding acceptance of the null hypothesis. But certainly, any such attention effects would appear to be vanishingly small, since Experiment 2 had sufficient power to detect attention effects one twelfth the size of those observed in Experiment 1. This suggests that differential visual attention to targets, as compared with distractors, is not necessary for MOT and we need to look elsewhere for the primary mechanism underlying object tracking.

One possibility is that object tracking uses a mechanism like that proposed by Pylyshyn (2003) in his visual indexing theory. Pointers are assigned to tracked objects at the beginning of the trial and continue to be bound to their indexed objects throughout tracking. According to Pylyshyn (2003), object indexes do not contain information about properties of the object and so would be unaffected by whether the object briefly increased in luminance. The index itself does not even contain explicit information about the location of the tracked object. It just provides a pointer that can be used to access properties of the indexed object when needed. In fact, according to index theory, there should not necessarily be any advantage to detecting probe flashes on tracked objects versus distractors, because the index is attached to the tracked object and should not provide privileged access to nearby probes (which constitute separate objects).

Current evidence suggests that an indexing mechanism with properties like these is located in the inferior intraparietal sulcus. Recent fMRI experiments show that this area is activated during MOT, with activation levels increasing with target set size (Culham et al., 1998). Xu and Chun (2009) reviewed evidence that this area is capable of individuating a maximum of four objects during working memory tasks, which is similar to the maximum number of tracked objects in MOT suggested by Pylyshyn (2001; Pylyshyn & Storm, 1988). Interestingly, activation in this area appears to depend only on the number of objects and not on their complexity (Xu & Chun, 2006), which is consistent with Pylyshyn's (e.g., 2001, 2003) suggestion that indexes are merely pointers to objects and do not themselves carry information about the features or properties of indexed objects.

If tracking is carried out primarily by a visual indexing mechanism, visual attention may play an important

ancillary role in reducing interference from distractors. An important source of errors during MOT is the tendency to mistake nearby distractors for targets, resulting in tracking of the distractor (O'Hearn et al., 2005; Pylyshyn, 2004). This may be viewed as a failure to separately individuate the target and distractor and may reflect the coarse spatial resolution of visual representations in the parietal area (Intriligator & Cavanagh, 2001). Visual attention is known to increase spatial resolution (Yeshurun & Carrasco, 1999) and could reduce errors in tracking by improving target individuation. This could take the form of target enhancement, distractor inhibition, or both. This error reduction process would clearly become more important as display density increases, and this may explain why tracking was associated with visual attention to targets in Experiment 3, but not in Experiment 2.

Enhancement Versus Suppression

Are the attention effects that we observed due to enhancement of the target, suppression of the distractor, or some combination of the two? We will consider this question for two separate sets of measures: early (N1 amplitude) and late (behavior and P300). In Experiment 1, we found evidence for both target enhancement and distractor suppression in the anterior N1 component, whereas the posterior N1 showed a "pure" distractor suppression pattern. Experiment 3 showed "pure" target enhancement in the anterior N1 and a distractor suppression pattern for the posterior N1 (although statistical support for the latter effect was weak). These results suggest that both target enhancement and distractor suppression may be present at early stages in visual processing of objects during MOT and that these effects may be mediated by visual attention.

These early effects of attention, however, appear to be dissociable from later attention effects reflected in probe detection accuracy and P300 amplitude. In Experiment 2, we found that N1 amplitude was the same for all probe locations, whereas detection accuracy and P300 amplitude showed strong effects of probe location (background > target > distractor). Pylyshyn and colleagues (Pylyshyn, 2006; Pylyshyn et al., 2008) have used detection performance for probes presented during MOT to argue that distractors are suppressed, particularly when they are likely to be confused with targets. This conclusion rests on the finding that distractor probes are detected less often than probes on targets and blank background areas (Pylyshyn, 2006). However, many (but not all) experiments have also shown that accuracy for background probes was superior to that for target probes, a pattern that we observed as well in Experiments 2 and 3. This raises the question of whether background probes may simply produce stronger signals than do object probes purely due to their physical properties, such as an absence of masking by nearby contours. If this were the case, it would invalidate using background probes as a *neutral condition* for inferring enhancement and suppression effects.

Using texture-defined objects like those employed here, Flombaum et al. (2008) continued to find an advantage for background probes, which argues against a simple (luminance) contour interference explanation. In a subsequent

study, Flombaum and Scholl (2009) carried out a series of experiments investigating the background probe advantage, including an experiment in which the target and distractor objects occasionally stopped during the trial. Detection accuracy for probes on stopped objects was much higher, relative to probes on these same objects when they were moving. This finding implicates movement as the source of poorer performance on object probes, as compared with background probes, which are generally stationary. This was confirmed in an experiment using texture-defined objects presented on either stationary or moving texture backgrounds. They found that presenting probes on a moving background eliminated their detection advantage and reduced performance to levels that were slightly worse than that for nearby targets and distractors. The reason that moving objects impair probe detection may be that the same transient channels that are responsible for detecting the onset of the probe are also involved in processing motion of the moving object (Breitmeyer, 1984).

Note that, in our experiment, moving backgrounds did not eliminate the background probe advantage, which might appear to be a failure to replicate the Flombaum and Scholl (2009) results. However, their probes were small squares that might have been effectively masked by the moving pixels of the background. Our probes were much larger and might not be masked by the moving background, which was composed of small micropatterns of random black and white pixels. In any case, these results together suggest that the background probe advantage in detection may be due, at least in part, to *greater masking of probes by moving objects*, as compared with probes on stationary backgrounds. The question remains whether there is any evidence of distractor suppression after eliminating the differential effects of masking.

Plyshyn et al. (2008) recently reported an experiment in which half of the distractors appeared on a different depth plane than did the target. They found that target and different-depth distractors were detected equally well and better than same-depth distractors, suggesting that distractors that can be preattentively grouped away from targets are not inhibited. This method provides evidence for distractor suppression without the confound of background probes appearing on stationary backgrounds.

In summary, it appears that the pattern of probe detection accuracy that we and others have observed (background > target > distractor) may be due, at least in part, to a stronger signal associated with background probes that are not masked by a moving object, as are the object probes. This possibility is reinforced by our finding that the amplitude of the P2 component of the ERP was enhanced for background probes, relative to both targets and distractors, which were similar to each other. Previous research (Luck & Hillyard, 1994a) showed that popout targets presented during a visual search task were associated with larger P2 components, suggesting that our background probes may have had similar popout properties. The ability of background probes to automatically attract attention might help explain why they are detected more often than probes appearing on objects. This suggests that background probes may not constitute a valid

neutral condition for inferring enhancement versus suppression effects and that other procedures, such as that used by Pylyshyn et al. (2008), may be required to make this determination.

The fact that the background probe advantage is observed in a component (the P2) that occurs later than the N1 is consistent with our observation that the detection advantage for background probes and the large P2 that accompanies this advantage can be observed in circumstances in which earlier components, such as the N1, show no effect whatsoever of probe location (Experiment 2). This suggests that the reduced masking associated with background probes affects processes that occur after the N1 and that the pattern of distractor suppression that we observed in the N1 component in Experiments 1 and 2 may not be affected by the masking confound. This is also consistent with the fact that the background N1 is never larger than the target N1, as it is in the case of the late-occurring P300.

Supporting evidence can also be found in the results of Drew et al. (2009), who examined three different nontarget probes: moving distractor, stationary distractor, and empty space. If the sort of differential masking effects outlined above affected N1 amplitude, we would expect small N1s for the moving object distractor, relative to both the stationary distractor and empty space conditions. However, contrary to the masking predictions, N1 amplitude was slightly (but significantly) larger for the moving distractor probes than for both of the other two probe types. Thus, there appears to be little evidence of a background probe advantage appearing in the N1 component. Nonetheless, it will be important to replicate our N1 suppression pattern using conditions like those employed by Pylyshyn et al. (2008), in which all probes occur on moving objects. On the basis of this logic, Doran and Hoffman (2009) carried out an experiment that examined whether distractor suppression depended on the distance of distractors from the tracked objects. They found a distinct distractor suppression pattern in which probes appearing on distractors near to the target object elicited smaller posterior N1s than did probes appearing on far distractors. Target and far distractor probes were equivalent.

Regardless of whether the late measures reflect target enhancement or distractor suppression, an intriguing question remains regarding the dissociation of attention effects between early and late measures observed in Experiment 2. Detection accuracy and P300 amplitude were larger for targets than for distractors even though N1 amplitudes were the same. We used this dissociation above to argue that an indexing mechanism, rather than visual attention, may be the primary mechanism involved in tracking. How could an indexing mechanism produce differences in probe detection without affecting N1 amplitude? One possibility is that indexes result in a *sustained activation* of target object representations in object-related areas of the visual cortex such as the lateral occipital complex (LOC). These target object activations could, in turn, suppress distractor object representations through the kind of winner-take-all network proposed by Desimone and Duncan (1995). The LOC is an important area for processing

objects (Grill-Spector, Kourtzi, & Kanwisher, 2001) and also appears to be the generator site of the posterior N1 (Martínez, Ramanathan, et al., 2007).

Sustained differences between targets and distractors would essentially show up as differences in baseline amplitude in ERPs and would be eliminated by the usual process of baseline correction. In other words, differences in baseline activation levels between targets and distractors would not appear in the amplitude of the N1 component elicited by the probes. These differences in activation could, however, affect how quickly and accurately information about probe flashes could be accessed. It is interesting to note that Drew and Vogel (2008) described an ERP component (the contralateral delay activity) that is sustained throughout the tracking trial and whose amplitude increases with target set size. This component appears to have a scalp distribution that is similar to the N1 and may reflect the sustained activation of target representations in the LOC proposed above.

What implications does the preceding discussion have for the N1 suppression effects we observed in Experiment 1? Clearly, it suggests caution in interpreting the pattern of N1 amplitudes in Experiment 1 as evidence of distractor suppression. Additional research using better neutral conditions, such as those described in Pylyshyn et al. (2008), needs to be combined with ERP measures to examine whether N1 amplitude reflects target enhancement or distractor suppression. The important point for the present is that there is clear evidence that visual attention plays a role in MOT, particularly for dense displays.

Conclusion

Our results, based on ERPs elicited by probe flashes on targets, distractors, and background areas, show that observers preferentially allocate visual attention to targets during MOT, a conclusion in agreement with Drew et al. (2009). However, we also showed that when tracking load is light and probes are task relevant, observers are able to distribute their attention broadly over the display without much loss in tracking accuracy, suggesting that visual attention may not be the *primary mechanism* underlying MOT. Object tracking may rely on an *indexing* mechanism that is distinct from visual attention. Increasing the difficulty of the tracking task resulted in a reinstatement of visual attention effects and a loss in probe detection accuracy. These results suggest that visual attention may be particularly important during tracking of objects in dense displays where distractors may be confused with targets.

AUTHOR NOTE

This work was supported in part by NIH Grant RO1 NS 050876, Subaward 8508-53692, to J.E.H. from Johns Hopkins University. Thanks to Todd Horowitz, Ed Vogel, Jeremy Wolfe, and an anonymous reviewer for helpful comments on an earlier draft. Thanks to Meg Lamborn, David McCoy, Hyun Young Park, Amanda Skoranski, and Laura Yarnall for their invaluable assistance in collecting data. Correspondence concerning this article should be addressed to M. M. Doran, Department of Psychology, University of Delaware, Newark, DE 19716 (e-mail: mdoran@psych.udel.edu).

REFERENCES

- ALVAREZ, G. A., & FRANCONERI, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, **7**(13, Art. 14), 1-10.
- ALVAREZ, G. A., HOROWITZ, T. S., ARSENIO, H. C., DiMASE, J. S., & WOLFE, J. M. (2005). Do multielement visual tracking and visual search draw continuously on the same visual attention resources? *Journal of Experimental Psychology: Human Perception & Performance*, **31**, 643-667.
- AVRBACH, E., & CORIELL, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, **40**, 309-328.
- AWH, E., & PASHLER, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 834-846.
- BACON, W. F., & EGETH, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, **55**, 485-496.
- BLASER, E., PYLYSHYN, Z. W., & HOLCOMBE, A. O. (2000). Tracking an object through feature space. *Nature*, **408**, 196-199.
- BREitmeyer, B. G. (1984). *Visual masking: An integrative approach*. New York: Oxford University Press.
- CARDINAL, R. N., & AITKEN, M. R. F. (2006). *ANOVA for the behavioral sciences researcher*. Mahwah, NJ: Erlbaum.
- CAVANAGH, P., & ALVAREZ, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, **9**, 349-354.
- COULL, J. T. (1998). Neural correlates of attention and arousal: Insights from electrophysiology, functional neuroimaging and psychopharmacology. *Progress in Neurobiology*, **55**, 343-361.
- COWAN, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral & Brain Sciences*, **24**, 87-185.
- CULHAM, J. C., BRANDT, S. A., CAVANAGH, P., KANWISHER, N. G., DALE, A. M., & TOOTELL, R. B. H. (1998). Cortical fMRI activation produced by attentive tracking of moving targets. *Journal of Neurophysiology*, **80**, 2657-2670.
- DESIMONE, R., & DUNCAN, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, **18**, 193-222.
- DI RUSSO, F., MARTÍNEZ, A., & HILLYARD, S. A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cerebral Cortex*, **13**, 486-499.
- DI RUSSO, F., MARTÍNEZ, A., SERENO, M. I., PITZALIS, S., & HILLYARD, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, **15**, 95-111.
- DONCHIN, E. (1981). Surprise! . . . Surprise? *Psychophysiology*, **18**, 493-513.
- DONCHIN, E., & COLES, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral & Brain Sciences*, **11**, 357-374.
- DORAN, M. M., & HOFFMAN, J. E. (2009). *Early distractor suppression in multiple object tracking*. Manuscript in preparation.
- DREW, T., MCCOLLOUGH, A. W., HOROWITZ, T. S., & VOGEL, E. K. (2009). Attentional enhancement during multiple-object tracking. *Psychonomic Bulletin & Review*, **16**, 411-417.
- DREW, T., & VOGEL, E. K. (2008). Neural measures of individual differences in selecting and tracking multiple moving objects. *Journal of Neuroscience*, **28**, 4183-4191.
- DUNCAN, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, **113**, 501-517.
- DUNCAN-JOHNSON, C. C., & DONCHIN, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, **14**, 456-467.
- EGLY, R., DRIVER, J., & RAFAL, R. D. (1994). Shifting visual-attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, **123**, 161-177.
- EIMER, M. (1997). Uninformative symbolic cues may bias visual-spatial attention: Behavioral and electrophysiological evidence. *Biological Psychology*, **46**, 67-71.
- ERIKSEN, B. A., & ERIKSEN, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, **16**, 143-149.

- ERIKSEN, C. W., & HOFFMAN, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics*, **12**, 201-204.
- ERIKSEN, C. W., & HOFFMAN, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perception & Psychophysics*, **14**, 155-160.
- ERIKSEN, C. W., & ST. JAMES, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, **40**, 225-240.
- ERIKSEN, C. W., & YEH, Y. Y. (1985). Allocation of attention in the visual-field. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 583-597.
- FAUL, F., ERDFELDER, E., LANG, A.-G., & BUCHNER, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, **39**, 175-191.
- FLOMBAUM, J. I., & SCHOLL, B. J. (2009). *The influence of motion on attention*. Manuscript under revision.
- FLOMBAUM, J. I., SCHOLL, B. J., & PYLYSHYN, Z. W. (2008). Attentional resources in visual tracking through occlusion: The high-beams effect. *Cognition*, **107**, 904-931.
- FRANCONERI, S. L., ALVAREZ, G. A., & ENNS, J. T. (2007). How many locations can be selected at once? *Journal of Experimental Psychology: Human Perception & Performance*, **33**, 1003-1012.
- FRANCONERI, S. L., LIN, J. Y., PYLYSHYN, Z. W., FISHER, B., & ENNS, J. T. (2008). Evidence against a speed limit in multiple-object tracking. *Psychonomic Bulletin & Review*, **15**, 802-808.
- GRILL-SPECTOR, K., KOURTZI, Z., & KANWISHER, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, **41**, 1409-1422.
- HELMHOLTZ, H. von (1925). *Treatise on physiological optics* (Translated from 3rd German ed.). New York: Dover. (Original work published 1867)
- HILLYARD, S. A., & ANLLO-VENTO, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences*, **95**, 781-787.
- HILLYARD, S. A., & MÜNTE, T. F. (1984). Selective attention to color and location: An analysis with event-related brain potentials. *Perception & Psychophysics*, **36**, 185-198.
- HOFFMAN, J. E., & NELSON, B. (1981). Spatial selectivity in visual search. *Perception & Psychophysics*, **30**, 283-290.
- HOROWITZ, T. S., BIRNKRAUT, R. S., FENCSIK, D. E., TRAN, L., & WOLFE, J. M. (2006). How do we track invisible objects? *Psychonomic Bulletin & Review*, **13**, 516-523.
- HULLEMAN, J. (2005). The mathematics of multiple object tracking: From proportions correct to number of objects tracked. *Vision Research*, **45**, 2298-2309.
- IMAI, A., & TSUJI, K. (2004). Event-related potential correlates of judgment categories and detection sensitivity in a visual detection task. *Vision Research*, **44**, 763-773.
- INTRILIGATOR, J., & CAVANAGH, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, **43**, 171-216.
- KUTAS, M., McCARTHY, G., & DONCHIN, E. (1977). Augmenting mental chronometry: P300 as a measure of stimulus evaluation time. *Science*, **197**, 792-795.
- LAVIE, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences*, **9**, 75-82.
- LAVIE, N., HIRST, A., DE FOCKERT, J. W., & VIDING, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, **133**, 339-354.
- LUCK, S. J., FAN, S., & HILLYARD, S. A. (1993). Attention-related modulation of sensory-evoked brain activity in a visual-search task. *Journal of Cognitive Neuroscience*, **5**, 188-195.
- LUCK, S. J., & HILLYARD, S. A. (1994a). Electrophysiological correlates of feature analysis during visual-search. *Psychophysiology*, **31**, 291-308.
- LUCK, S. J., & HILLYARD, S. A. (1994b). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 1000-1014.
- LUCK, S. J., & VOGEL, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.
- LUCK, S. J., WOODMAN, G. F., & VOGEL, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, **4**, 433-440.
- MALINOWSKI, P., FUCHS, S., & MÜLLER, M. M. (2007). Sustained division of spatial attention to multiple locations within one hemifield. *Neuroscience Letters*, **414**, 65-70.
- MANDLER, G., & SHEBO, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, **111**, 1-22.
- MANGUN, G. R., & HILLYARD, S. A. (1988). Spatial gradients of visual-attention: Behavioral and electrophysiological evidence. *Electroencephalography & Clinical Neurophysiology*, **70**, 417-428.
- MANGUN, G. R., & HILLYARD, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual spatial priming. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 1057-1074.
- MARTÍNEZ, A., DI RUSSO, F., ANLLO-VENTO, L., SERENO, M. I., BUXTON, R. B., & HILLYARD, S. A. (2001). Putting spatial attention on the map: Timing and localization of stimulus selection processes in striate and extrastriate visual areas. *Vision Research*, **41**, 1437-1457.
- MARTÍNEZ, A., RAMANATHAN, D. S., FOXE, J. J., JAVITT, D. C., & HILLYARD, S. A. (2007). The role of spatial attention in the selection of real and illusory objects. *Journal of Neuroscience*, **27**, 7963-7973.
- MARTÍNEZ, A., TEDER-SÄLEJÄRVI, W., & HILLYARD, S. A. (2007). Spatial attention facilitates selection of illusory objects: Evidence from event-related brain potentials. *Brain Research*, **1139**, 143-152.
- MARTÍNEZ, A., TEDER-SÄLEJÄRVI, W., VAZQUEZ, M., MOLHOLM, S., FOXE, J. J., JAVITT, D. C., ET AL. (2006). Objects are highlighted by spatial attention. *Journal of Cognitive Neuroscience*, **18**, 298-310.
- MCMAINS, S. A., & SOMERS, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron*, **42**, 677-686.
- MOST, S. B., SCHOLL, B. J., CLIFFORD, E. R., & SIMONS, D. J. (2005). What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review*, **112**, 217-242.
- MÜLLER, M. M., MALINOWSKI, P., GRUBER, T., & HILLYARD, S. A. (2003). Sustained division of the attentional spotlight. *Nature*, **424**, 309-312.
- O'CRAVEN, K. M., DOWNING, P. E., & KANWISHER, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature*, **401**, 584-587.
- O'HEARN, K., LANDAU, B., & HOFFMAN, J. E. (2005). Multiple object tracking in people with Williams syndrome and in normally developing children. *Psychological Science*, **16**, 905-912.
- PASHLER, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, **44**, 369-378.
- POLICH, J., & KOK, A. (1995). Cognitive and biological determinants of P300: An integrative review. *Biological Psychology*, **41**, 103-146.
- POSNER, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, **32**, 3-25.
- POSNER, M. I., SNYDER, C. R. R., & DAVIDSON, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, **109**, 160-174.
- PYLYSHYN, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, **80**, 127-158.
- PYLYSHYN, Z. W. (2003). *Seeing and visualizing: It's not what you think*. Cambridge, MA: MIT Press, Bradford Books.
- PYLYSHYN, Z. W. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition*, **11**, 801-822.
- PYLYSHYN, Z. W. (2006). Some puzzling findings in multiple object tracking (MOT): II. Inhibition of moving nontargets. *Visual Cognition*, **14**, 175-198.
- PYLYSHYN, Z. W., & ANNAN, V. (2006). Dynamics of target selection in multiple object tracking (MOT). *Spatial Vision*, **19**, 485-504.
- PYLYSHYN, Z. W., HALADJIAN, H., KING, C., & REILLY, J. (2008). Selective nontarget inhibition in multiple object tracking (MOT). *Visual Cognition*, **16**, 1011-1021.
- PYLYSHYN, Z. W., & STORM, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, **3**, 179-197.
- SCHOLL, B. J. (2001). Objects and attention: The state of the art. *Cognition*, **80**, 1-46.
- SCHOLL, B. J. (2009). What have we learned about attention from mul-

- multiple object tracking (and vice versa)? In D. Dedrick & L. Trick (Eds.), *Computation, cognition, and Pylyshyn* (pp. 49-77). Cambridge, MA: MIT Press.
- SCHOLL, B. J., & PYLYSHYN, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, **38**, 259-290.
- SCHOLL, B. J., PYLYSHYN, Z. W., & FELDMAN, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*, **80**, 159-177.
- SEARS, C. R., & PYLYSHYN, Z. W. (2000). Multiple object tracking and attentional processing. *Canadian Journal of Experimental Psychology*, **54**, 1-14.
- SHIM, W. M., ALVAREZ, G. A., & JIANG, Y. V. (2008). Spatial separation between targets constrains maintenance of attention on multiple objects. *Psychonomic Bulletin & Review*, **15**, 390-397.
- SIBLY, M. (2005). Blitz3D (Version 1.90) [Computer software]. Auckland, New Zealand: Blitz Research Ltd.
- SPERLING, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74**, 1-29.
- TRICK, L. M., & PYLYSHYN, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, **101**, 80-102.
- VALDES-SOSA, M., BOBES, M. A., RODRIGUEZ, V., & PINILLA, T. (1998). Switching attention without shifting the spotlight: Object-based attentional modulation of brain potentials. *Journal of Cognitive Neuroscience*, **10**, 137-151.
- VANMARLE, K., & SCHOLL, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, **14**, 498-504.
- VOGEL, E. K., WOODMAN, G. F., & LUCK, S. J. (2005). Pushing around the locus of selection: Evidence for the flexible-selection hypothesis. *Journal of Cognitive Neuroscience*, **17**, 1907-1922.
- WRIGHT, R. D., & WARD, L. M. (2008). *Orienting of attention*. New York: Oxford University Press.
- XU, Y. D., & CHUN, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, **440**, 91-95.
- XU, Y. D., & CHUN, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, **13**, 167-174.
- YESHURUN, Y., & CARRASCO, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research*, **39**, 293-306.

NOTES

1. It can be difficult to compare accuracy across MOT experiments that use different set sizes, since changes in the number of targets and/or the number of distractors affect guessing probabilities. However, it is possible to correct for differences in guessing probabilities, using formulas such as the one provided by Pylyshyn and Annan (2006; see also Hulleman, 2005). In essence, these formulas take the number of objects correctly tracked on each trial and remove the probability of guessing correctly to create a measure of the number of objects effectively tracked.

2. Note that the use of multiple follow-up *t* tests is appropriate when an ANOVA produces a significant main effect of a factor with three levels. In this case, there is no need to use any correction procedure (e.g., Bonferroni correction), since the family-wise Type I error rate does not become inflated (Cardinal & Aitken, 2006).

3. In addition, this sort of attention switching seems unlikely, given that the same sort of attentional set would be useful in the following Experiment 3, but once again, attentional selection was observed in the N1.

(Manuscript received December 22, 2008;
revision accepted for publication July 19, 2009.)