

# Spatio-temporal Patterns of Brain Activity Distinguish Strategies of Multiple-object Tracking

Christian Merkel<sup>1</sup>, Christian M. Stoppel<sup>1</sup>, Steven A. Hillyard<sup>2,3</sup>,  
Hans-Jochen Heinze<sup>1,2</sup>, Jens-Max Hopf<sup>1,2</sup>,  
and Mircea Ariel Schoenfeld<sup>1,2,4</sup>

## Abstract

Human observers can readily track up to four independently moving items simultaneously, even in the presence of moving distractors. Here we combined EEG and magnetoencephalography recordings to investigate the neural processes underlying this remarkable capability. Participants were instructed to track four of eight independently moving items for 3 sec. When the movement ceased a probe stimulus consisting of four items with a higher luminance was presented. The location of the probe items could correspond fully, partly, or not at all with the tracked items. Participants reported whether the probe items fully matched the tracked items or not. About half of the participants showed slower RTs and higher error rates with increasing correspondence between tracked items and the probe. The other half, however, showed faster RTs and lower error rates when

the probe fully matched the tracked items. This latter behavioral pattern was associated with enhanced probe-evoked neural activity that was localized to the lateral occipital cortex in the time range 170–210 msec. This enhanced response in the object-selective lateral occipital cortex suggested that these participants performed the tracking task by visualizing the overall shape configuration defined by the vertices of the tracked items, thereby producing a behavioral advantage on full-match trials. In a later time range (270–310 msec) probe-evoked neural activity increased monotonically as a function of decreasing target–probe correspondence in all participants. This later modulation, localized to superior parietal cortex, was proposed to reflect the degree of mismatch between the probe and the automatically formed visual STM representation of the tracked items. ■

## INTRODUCTION

In everyday life, the amount of information impinging on the eyes far exceeds the processing capacities of the visual system. Therefore, mechanisms that separate relevant from irrelevant information are critical to adaptive behavior. The most prominent and arguably the most effective selection mechanism is based on spatial location and has been likened to a spotlight that facilitates processing within a selected location (Posner, 1980) while inhibiting adjacent inputs (Hopf et al., 2006). It is also well established that attention can be directed to unitary objects, thereby facilitating the processing of all features that belong to the selected object (Martinez, Ramanathan, Foxe, Javitt, & Hillyard, 2007; Schoenfeld, Tempelmann, et al., 2003; O'Craven, Downing, & Kanwisher, 1999; Duncan, 1984). Both behavioral and physiological studies have shown that the attentional spotlight can be divided such that more than one spatial location can be selected for facilitated processing (Dubois, Hamker, & VanRullen, 2009; McMains & Somers, 2004; Muller, Malinowski,

Gruber, & Hillyard, 2003). Although a divided spotlight mechanism may be effective for attending to two different objects at the same time, it faces a serious challenge when the situation requires simultaneous processing of a higher number of objects at different locations. This leads to the question how effectively the visual system is able to attend selectively to multiple objects in the presence of distractors.

Pylyshyn and colleagues investigated this question by asking participants to keep track of a subset of visually identical objects moving randomly about in the visual field (Pylyshyn, 1989, 2004; Pylyshyn & Storm, 1988). They found that participants' performance decreased in proportion to the number of objects to be tracked. Surprisingly, however, the majority of participants were able to keep track at least four objects simultaneously (Alvarez & Franconeri, 2007; Pylyshyn & Storm, 1988), a feat that appears to exceed the capacity limitations assumed by various theories of attention (Schneider, 1993; Eriksen & St James, 1986; Duncan, 1984; Egeth, Virzi, & Garbart, 1984; Treisman & Schmidt, 1982; Posner, 1980). To account for this multiple-object tracking (MOT) capability, Pylyshyn and colleagues postulated a process that maintains the spatial information of the tracked items through time in a way that is independent of attentional

<sup>1</sup>Otto-von-Guericke University, Magdeburg, <sup>2</sup>Leibniz Institute for Neurobiology, Magdeburg, <sup>3</sup>University California, San Diego, <sup>4</sup>Kliniken Schmieder, Allensbach

requirements (FINST Model; Pylyshyn & Storm, 1988). However, some predictions of this model have not been fulfilled in subsequent experiments (Pylyshyn, 2004). Some studies have shown that the attentional spotlight can be divided, hence the processing in two different spatial locations can be concurrently enhanced. This could be first observed using EEG measures (Muller et al., 2003) but recently also with measures at single neuron level (Niebergall, Khayat, Treue, & Martinez-Trujillo, 2011). Other theoretical approaches have taken a multiple spotlights approach, in which the number of spotlights matches the number of targets to be tracked (Cavanagh & Alvarez, 2005).

A very different approach derived from object-based theories of attention has been proposed to account for MOT tracking performance. Scholl, Pylyshyn, and Feldman (2001) showed that the units of attentional selection during MOT are the target objects themselves. A further simplification of the problem of having multiple objects is the basic idea that the set of items to be tracked is preattentively grouped into a shape such as, for example, a polygon. Attentional processes are then directed toward this object, which continually changes shape as its vertices change location over time (Yantis, 1992). The integration of the multiple items into a single polygon helps to overcome the capacity limitations of selective attention, which then needs only to monitor a single object. A significant benefit of such perceptual grouping has been observed in other types of visual attention tasks. For example, the time to locate a target in visual search tasks is significantly reduced when stimulus properties allow for an integration of items into perceptual chunks (Treisman & Gormican, 1988; Treisman & Schmidt, 1982). These perceptual chunks are assembled on the basis of preattentive gestalt principles (Driver, Davis, Russell, Turatto, & Freeman, 2001; Rensink & Enns, 1995; Driver & Baylis, 1989), and spatio-temporal information about the whole configuration can be preferentially processed.

Previous paradigms have assessed the accuracy of MOT either by probing one of the several tracked targets (Alvarez & Franconeri, 2007; Yantis, 1992; Pylyshyn & Storm, 1988) or by asking participants to point toward the tracked items (Tombu & Seiffert, 2008; Scholl et al., 2001). Probing just one of the tracked items, however, does not reveal whether information about the overall spatial configuration of the items was being used to facilitate their tracking. To investigate such configurational processing, we designed a paradigm in which the probe consisted of multiple items that matched fully, partly or not at all with the tracked items. The hypothesis of configurational processing would be supported if participants responded faster and more accurately to probes that fully matched the array of tracked items.

At present only few studies investigated the electrophysiological correlates of neural processes involved in MOT (Drew, Horowitz, & Vogel, 2013; Drew, Horowitz, Wolfe, & Vogel, 2012; Drew, McCollough, Horowitz, &

Vogel, 2009). Typical changes in electrophysiological activity have been described during feature- (Bondarenko et al., 2012; Schoenfeld et al., 2007; Hillyard, Teder-Salejarvi, & Munte, 1998), space- (Boehler, Tsotsos, Schoenfeld, Heinze, & Hopf, 2009; Hopf et al., 2006; Mangun & Hillyard, 1988), and object-based (Schoenfeld, Woldorff, et al., 2003) selection. Here the mechanisms underlying item selection in the object-tracking task were investigated using ERPs and magnetic fields time-locked to the probe presented at the end of the tracking. The analyses will focus on the first event-related components exhibiting differences. Neural correlates of configurational processing would be expected to occur at an early time range and be associated with activity in the visual lateral occipital cortex (LO), the key region for object processing (Jiang et al., 2008; Schoenfeld, Woldorff, et al., 2003; Grill-Spector, Kushnir, Edelman, Itzhak, & Malach, 1998; Malach et al., 1995).

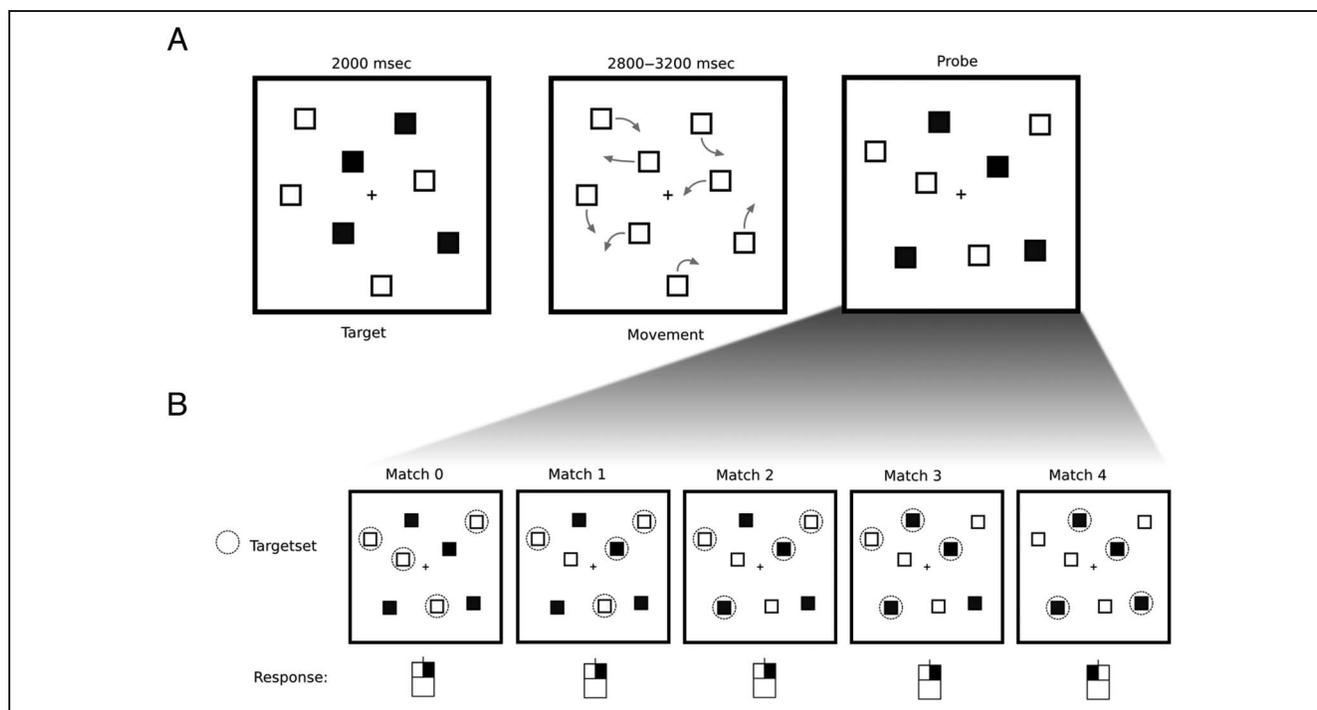
## METHODS

### Participants

Thirty-eight healthy participants (18 women, mean age = 25 years,  $SD = 2.9$  years) with normal or corrected-to-normal vision participated as paid volunteers. All participants gave informed consent, and the local ethics committee approved the study. None of the participants had performed a similar task before.

### Stimuli and Task

Stimuli were presented against a black background on a  $21^\circ \times 21^\circ$  video display (Figure 1A). A central fixation cross subtending  $0.5^\circ$  was present throughout the experiment. Each trial began with the appearance of eight identical items (white outline squares, each  $1^\circ \times 1^\circ$ ) randomly placed within the display. Four of those squares were flashed to a solid white three times within a 2-sec interval, thereby designating them as targets to be tracked. All eight squares then moved along randomized trajectories for 2800–3200 msec ( $180 \pm 12$  frames) until they reached final fixed positions that were identical across all trials and participants. (The movement trajectories for each trial were obtained by starting from the same spatial configuration and calculating randomly the movement of each target and distractor off-line for a certain number of steps. Then the sequence was reversed, in that each trial started with a different spatial configuration of targets and distractors but ended in the same spatial configuration across all trials.) Immediately following movement cessation, four of the eight squares became solid white again (the “probe” stimulus), and participants had to indicate by pushing one of two buttons whether or not all four of those probe items matched the previously tracked targets. Thus, participants pushed one button when either 0, 1, 2, or 3 of the probes matched the tracked items and the other button when all four matched (Figure 1B).



**Figure 1.** Experimental design. (A) At the beginning of each trial, four of the eight items were flashed to designate them as targets. All eight items then moved around randomly for about 3 sec, after which they stopped and four items were flashed as probes. The end positions and probes had the same spatial configuration on each trial. (B) The probe could be nonmatching, partly matching, or fully matching, depending on which targets were assigned at the beginning of each trial. Participants made one response to the full-match (match-4) probes and a different response to the nonmatching and partly matching probes.

Importantly, not only were the positions of all eight items in the final frame kept constant, but also the positions of the four probed items were identical on all trials. The degree of correspondence between tracked and probed items across conditions was thus varied by designating different subsets of the eight items as targets to be tracked at the beginning of each trial.

The movement trajectories of the eight items were calculated independently of each other to prevent participants from anticipating certain movement patterns such as “bouncing off” (Makovski, Vazquez, & Jiang, 2008; Scholl & Pylyshyn, 1999) and “force field” (Flombaum, Scholl, & Pylyshyn, 2008; Alvarez & Franconeri, 2007). Movement parameters were varied at random every 2–60 frames (33.3–1000 msec), resulting in a smooth unpredictable movement of each item. Constraints imposed during tracking ensured that (a) the minimal distance between the centers of two items never fell below  $1.8^\circ$ , (b) each item never touched the boundary of the  $21^\circ \times 21^\circ$  display, and (c) a trial never began with an item occluding the fixation cross. Each trajectory was used only once, so that each trial was composed of eight unique and independent trajectories.

## Procedure

Each participant performed in 300 trials overall, divided into 10 blocks that lasted 4–5 min each. Trials with each

of the five match conditions (60 trials each) were randomly distributed throughout the experiment. Participants were asked to track the four designated targets as they moved without moving their eyes. Upon the occurrence of the probe at the offset of the movement, participants were instructed to indicate as quickly as possible whether or not the probe matched all four of the tracked targets. On the “full-match trials” they pressed a button with the index finger, and otherwise they pressed another button with the middle finger.

## ERP and Event-related Magnetic Field Recording

The EEG was recorded from 32 electrodes (NeuroScan, Inc., El Paso, TX) placed according to the 10–20 system with a sampling rate of 509 Hz and a bandpass of DC 200 Hz. The horizontal EOG was recorded from a pair of electrodes placed on the left and right external canthi, and the vertical EOG was recorded from an electrode placed below the right eye. The right mastoid was used as reference during recording. Electrode impedances were kept below 5 k $\Omega$ . Magnetoencephalographic (MEG) activity was recorded concurrently with the EEG using a BTI Magnes 2500 WH (4-D Neuroimaging, San Diego, TX) whole head magnetometer containing 248 squids.

Artifact rejection was performed off-line using an individual peak-to-peak detection criterion for each participant that ranged between 80 and 130  $\mu$ V for the EEG and

2–3 pT for the MEG. For the electrophysiological data, an average of 13.69% of the trials were excluded from analysis. The amount of rejected trials did not differ across the different conditions,  $F(4, 148) = 0.47, p > .7$ . Furthermore, on average 15.82% of the MEG trials were discarded. Likewise there was no difference in the number of rejected trials across the conditions,  $F(4, 148) = 0.93, p > .4$ . ERP and event-related magnetic field (ERF) waveforms were averaged over 1000 msec epochs time-locked to the onset of the probe array, separately for each of the five target–probe matching conditions. The averaged ERPs were re-referenced to the average of the left and right mastoid before analysis.

## Data Analysis

### Behavior

Within-subject differences in error rate and accuracy were tested using repeated-measures ANOVA with the degree of probe–target match as a five-level factor. To further investigate between-subject differences in the use of global, configurational information to track multiple objects, participants were subdivided into two equally sized groups based on their error rates in the full-match probe condition relative to the zero-match condition, which served as a baseline. Additional repeated-measures ANOVAs were performed for each of these groups separately.

### ERP

ERPs elicited by the probes were quantified as a function of probe–target congruity level as mean amplitudes over two different time windows, 170–210 msec and 270–310 msec, relative to the mean baseline amplitude over a 200-msec prestimulus interval. ERPs were quantified at three occipital electrodes (O9, IZ, and O10), where the modulations associated with probe–target congruity were maximal. ERP amplitudes elicited by probes of differing congruity were compared using three-way repeated-measures ANOVAs with within-subject factors of congruity level and electrode/sensor site and group as a between-subject factor.

### Source Analyses

Source analyses of amplitude modulations as a function of probe–target congruity were carried out using multimodal neuroimaging software (Curry 6.0, Neuroscan Inc.). Source modeling was performed using minimum L2-norm estimates in a realistic boundary element model of the head derived from the MNI brain. For this analysis, the ERP and the simultaneously recorded ERF surface topographies were fit concurrently to obtain maximal localization power (Schoenfeld, Woldorff, et al., 2003; Fuchs et al., 1998).

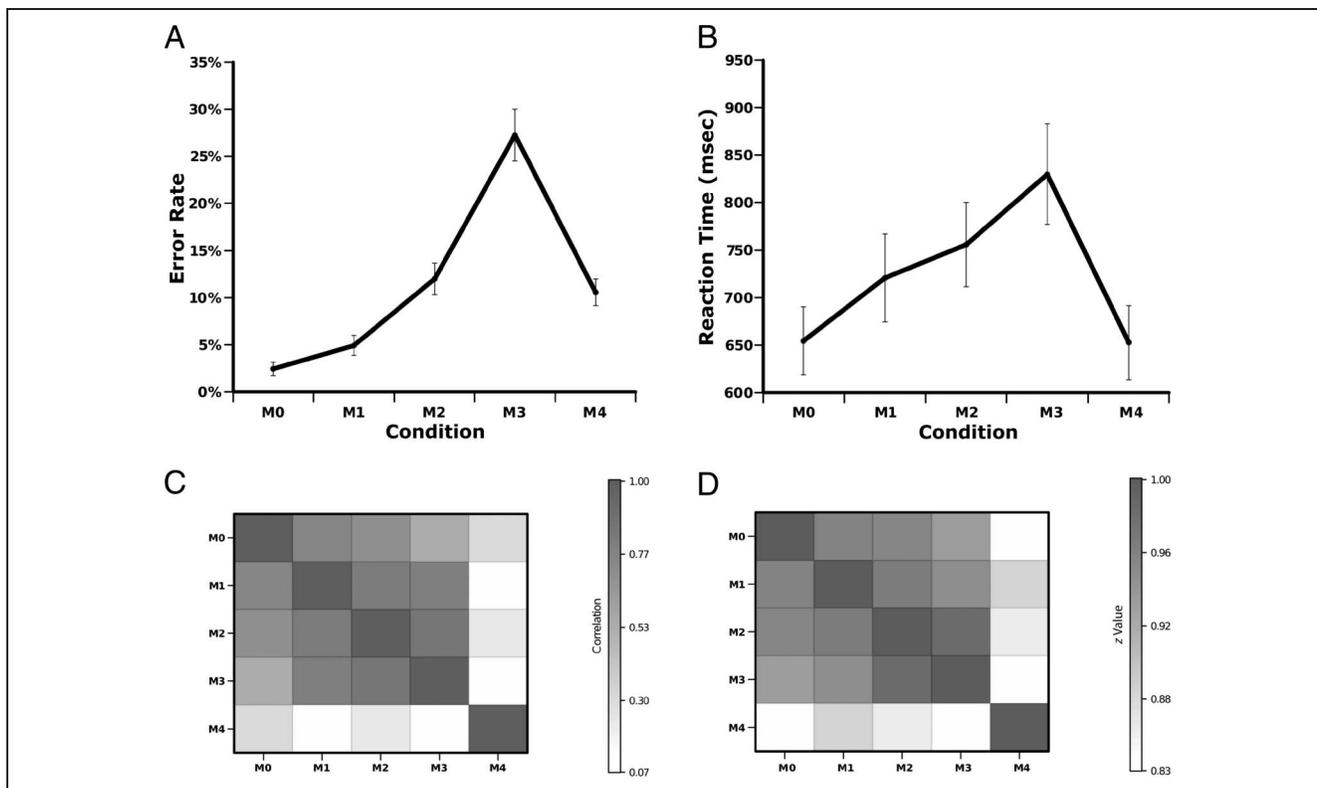
## RESULTS

### Behavior

Trials with RTs that deviated more than two standard deviations from the mean for each condition (6.5% of the trials) were considered as outliers and were excluded from further analysis. Error rates in this congruity judgment differed significantly as a function of the degree of matching between the probe and the tracked target items,  $F(4, 148) = 56.31, p < .001$  (Figure 2A). There was a progressive increase of error rate with increasing degree of congruity, except for the full-match (match-4) condition, for which significantly fewer errors were made in comparison with the match-3 condition,  $t(37) = -5.54, p < .001$ . RTs were similarly affected by the degree of probe–target congruity,  $F(4, 148) = 27.91, p < .001$  (Figure 2B), with faster RTs on full-match trials relative to match-3 trials,  $t(37) = 6.07, p < .001$ .

To investigate individual differences in patterns of responding, the participants' performance at each level of probe–target congruity was correlated with their performance at each of the other levels (Figure 2C and D). Compared with the correlations among match-0, match-1, match-2, and match-3 trials, the correlations across participants between the full-match (match-4) trials and any of the other trial types were considerably smaller. This was true for both error rates (Figure 2C) RTs (Figure 2D). These low correlations could be attributed to the wide variation in full-match performance among individual participants. Whereas all participants responded rapidly and with a low error rate on the match-0 trials, about half of the participants also responded rapidly and with a low error rate on the full-match trials. These latter participants also tended to make more errors on match-3 trials than did the other half of the participants. These distinctive performance patterns suggested that individual participants used different perceptual strategies when responding to fully matching versus partially matching probes. To further investigate these differences, participants were divided into two subgroups of equal size based on the difference in their error rates between the full-match and the match-0 conditions. Notably, this difference was smaller in Group 1 compared with Group 2. However, overall there were no differences for RT or error rates across groups. Two-way ANOVAs with factors Condition (M0, M1, M2, M3, M4) and Group (Group 1, Group 2) neither revealed a main effect for Group on the error rates,  $F(1, 36) = 0.018, p > .5$ , nor on the RTs,  $F(1, 36) = 0.038, p > .4$ .

Under regular circumstances splitting up a sample at its median might produce the potential problem of discarding relevant information by dichotomizing an in fact continuous variable (MacCallum, Zhang, Preacher, & Rucker, 2002; Cohen, 1983). However, if the variable to be down-sampled by the split maintains the natural clustering into discrete entities, the median split is feasible (MacCallum et al., 2002; Meehl, 1992). Following the median split, the variance explanation did not change for

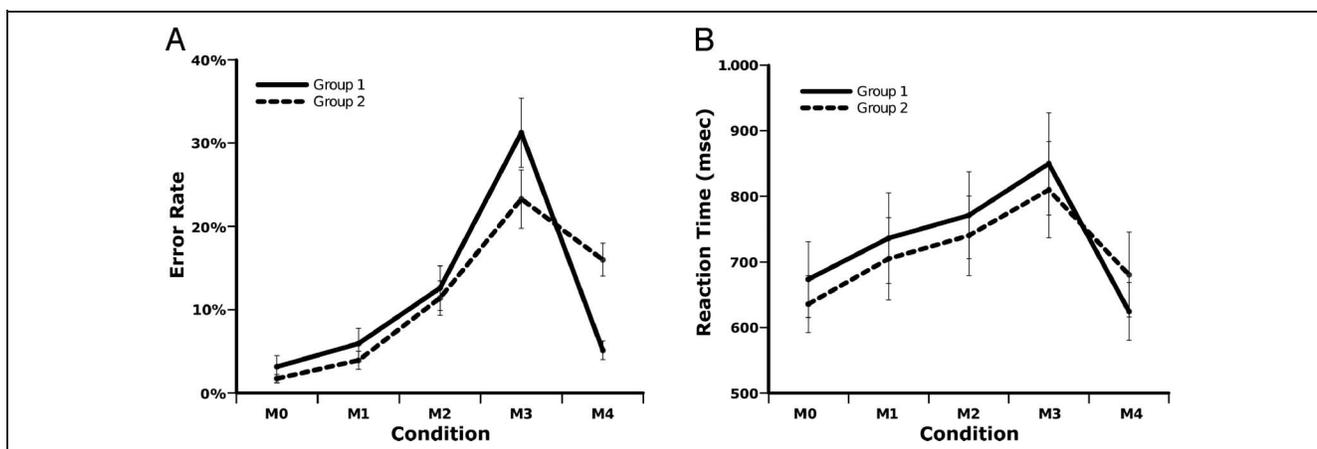


**Figure 2.** Behavioral results. Mean values over all participants of (A) error rates and (B) RTs as a function of match condition (probe–target congruity). (C) Pairwise correlations of error rates between the different match conditions across all participants. (D) Pairwise correlations of RTs between the different match conditions across all participants.

the continuous and dichotomized error rate difference. Thus, in this particular case, the between-group analysis of the sample split by its median error rate is, in this case, valid.

The main effect of matching condition (target–probe congruity) on the error rates remained significant within both groups [Group 1:  $F(4, 72) = 49.70, p < .001$ ; Group 2:  $F(4, 72) = 27.27, p < .001$ ; Figure 3A]. A two-way ANOVA with the factors Condition (five levels) and Group

(Group 1/Group 2) revealed an interaction between condition and group,  $F(4, 144) = 8.50, p < .001$ . The source of this interaction derived from the full-match condition, in which there were fewer errors in Group 1 than in Group 2,  $t(36) = 4.79, p < .001$ . Group 1 also showed much lower error rates on full-match trials than on match-3 trials,  $t(18) = -7.03, p < .001$ , whereas this comparison did not reach significance for Group 2,  $t(18) = -1.97, p > .05$ . The RTs showed a similar overall

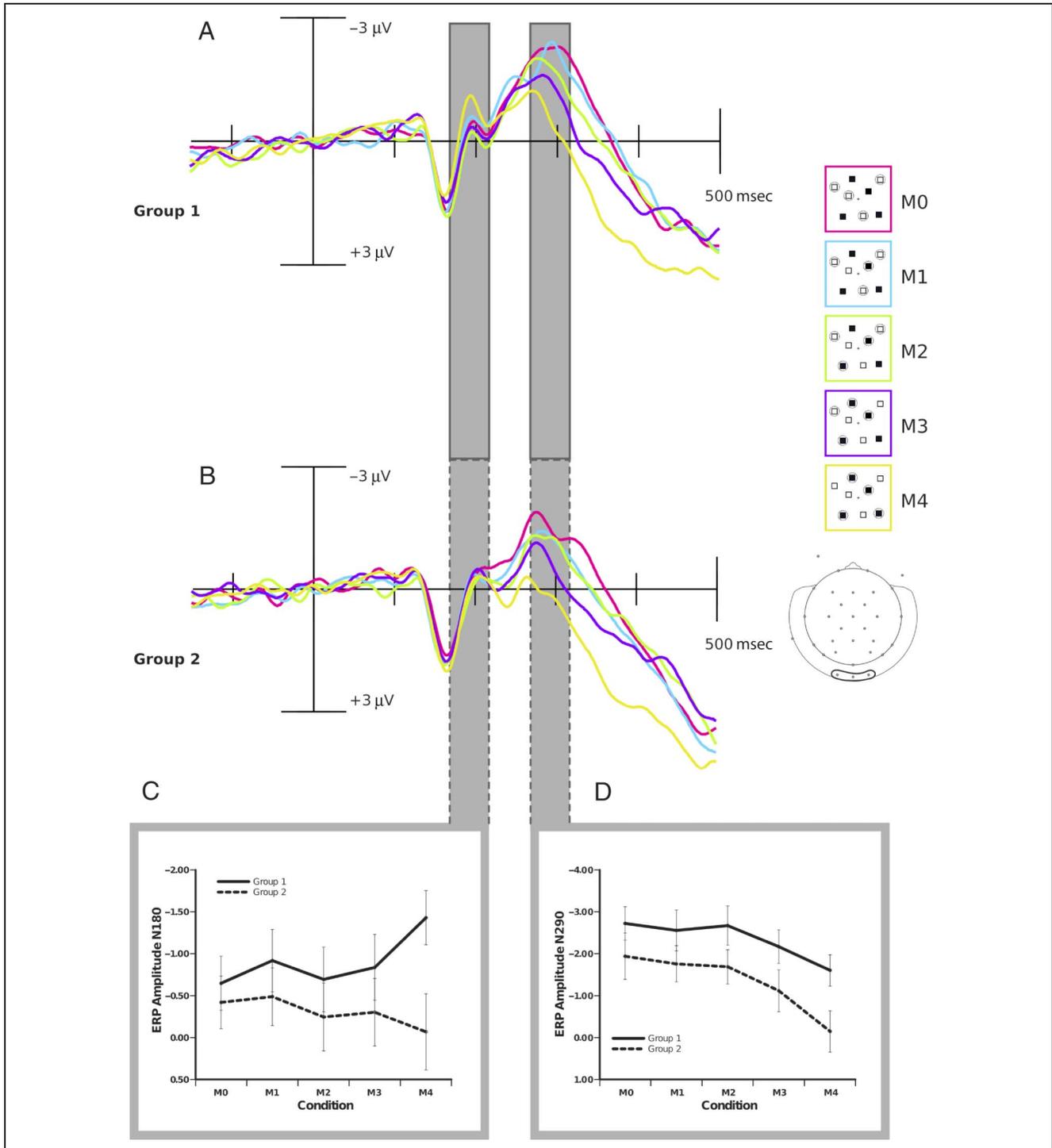


**Figure 3.** Comparisons between Group 1 and Group 2 participants in (A) error rate and (B) RT as a function of match condition (probe–target congruity). Note the differences between the groups in the match-3 and match-4 conditions.

pattern to error rates (Figure 3B), although the interaction between match condition and group only approached significance,  $F(4, 144) = 2.16, p < .08$ . Both groups showed a faster reaction toward match-4 trials compared with match-3 trials (Group 1:  $t(18) = -5.09, p < .001$ ; Group 2:  $t(18) = -3.60, p < .002$ ).

### ERP Data

The ERPs elicited by the probes showed target–probe congruity effects within the latency ranges 170–210 msec (the N180 component) and 270–310 msec (the N290 component; Figure 4). A three-way repeated-measures



**Figure 4.** ERPs to probes. (A) ERP waveforms for Group 1 averaged over sites O9, Iz, and O10 for the different matching conditions. Modulations in the N180 and N290 latency ranges are highlighted. (B) ERP waveforms for Group 2. (C) Mean N180 amplitude as a function of match condition. Note enlarged N180 on match-4 trials in Group 1 only. (D) Mean N290 amplitude as a function of match condition. Note monotonic decrease in N180 amplitude as a function of increasing probe–target congruity.

ANOVA for mean amplitude over the time range of 170–210 msec with the factors Group (Group 1, Group 2), condition (match-0 to match-4), and Electrode Site (O9, IZ, O10) revealed no significant main effect on N180 for Condition,  $F(4, 144) = 0.85, p > .4$ , or Group,  $F(1, 36) = 0.16, p > .2$ . There was, however, a significant interaction between the factors Group and Condition,  $F(4, 144) = 3.00, p < .02$ , which could be attributed to a larger N180 to the full-match probe in Group 1 than in Group 2,  $t(36) = 2.45, p < .02$  (Figure 4C). Moreover, N180 amplitude to the full-match probe was larger than to the match-3 probe in Group 1,  $t(18) = 2.29, p < .035$ , but not in Group 2,  $t(18) = -1.29, p > .2$ . This group difference was also evident in subsequent two-way ANOVAs carried out for each group separately: There was a significant main effect of Match Condition on N180 in Group 1,  $F(4, 72) = 3.80, p < .007$ , but not in Group 2,  $F(4, 72) = 0.68, p > .6$ .

Individual differences in the N180 elicited by the full-match probe were correlated with performance measures across participants. There was a significant correlation between the error rate difference and N180 amplitude difference between match-4 and match-0 trials,  $r(36) = -0.386, p < .018$ . A similar correlation was observed between the RT difference and the N180 amplitude difference between the match-4 and match-0 trials,  $r(36) = 0.412, p < .01$ . Thus, participants with larger N180 amplitudes to the match-4 probes tended to respond to those probes more accurately and rapidly than participants with lower N180 amplitudes.

For the mean ERP amplitude within the 270–310 msec interval (N290 component) a repeated-measures ANOVA with the factors Condition (match-0 through match-4) and Electrode Site (O9, IZ, O10) revealed significant effects for Condition in both Group 1,  $F(4, 72) = 4.84, p < .002$ , and Group 2,  $F(4, 72) = 10.61, p < .001$  (Figure 4D). A three-way ANOVA with factors of Condition, Electrode Site, and Group showed no interaction between Group and Condition,  $F(4, 144) = 0.78, p > .5$ . For both groups, the N290 showed a monotonic decrease in amplitude as a function of increasing congruity between target and probe.

### Source Analysis

Source analyses were carried out for the congruity-related modulations of the N180/M180 and N290/M290 components separately for Group 1 (Figure 5) and Group 2 (Figure 6). The ERP and ERF topographies were fit concurrently for the match-4 minus match-0 difference waveforms in the time ranges of these early and late modulations. In the N180/M180 time range, significant differences in amplitude between the match-4 and match-0 trials were only found for Group 1. Consequently, N180/M180 sources were modeled for Group 1 only, and the maximal current source densities were found to be located bilaterally in the middle occipital gyrus (BA 19, Talairach coordinates:

–50, –70, 0 for left hemisphere and 48, –77, 3 for right hemisphere) (Figure 5D).

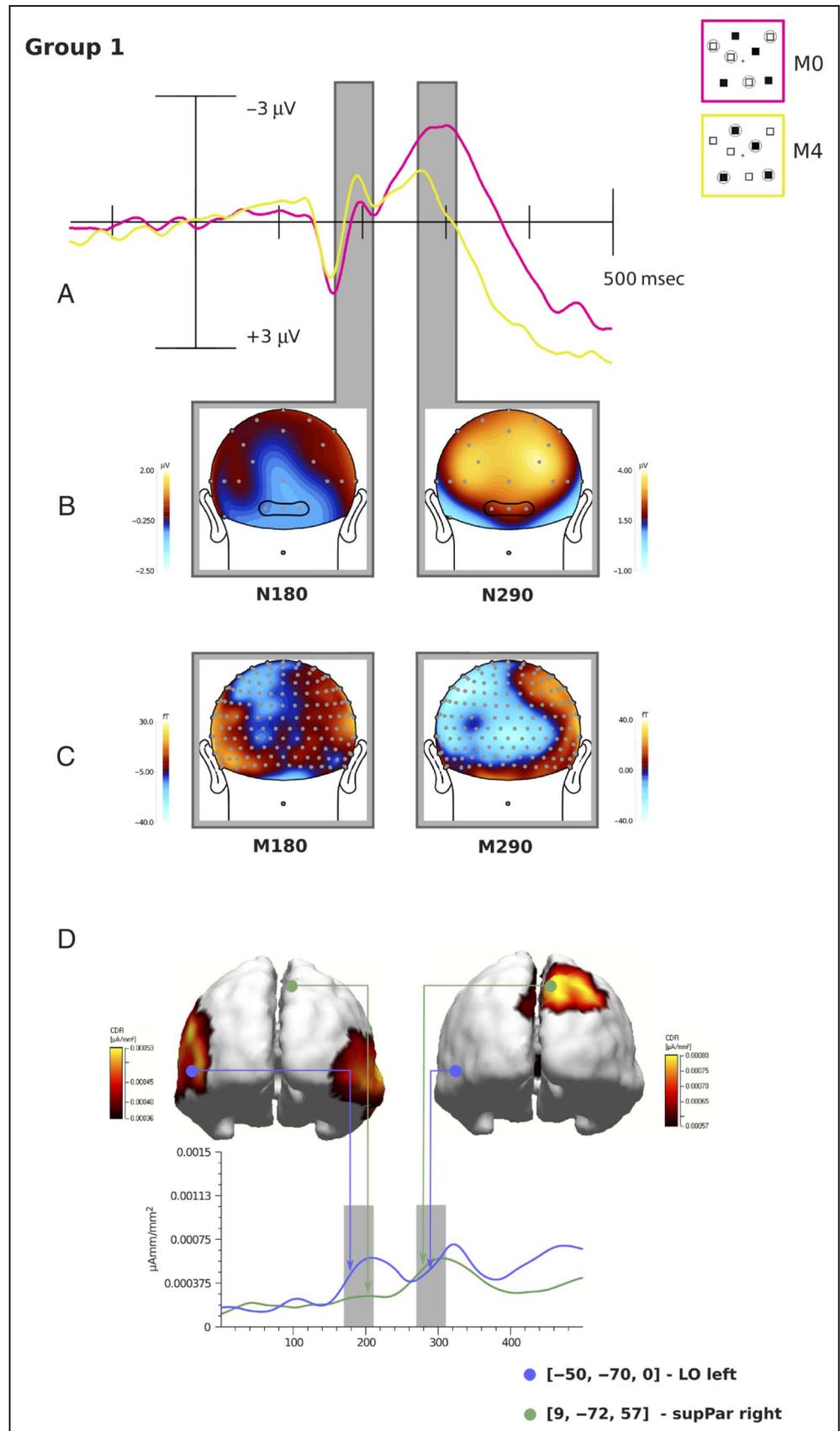
In the N290/M290 time range, significant amplitude differences between the match-4 and match-0 trials were observed in both groups. The modeled sources were located bilaterally in the superior parietal cortex (BA 7) in both groups, but the current density estimates were more extensive over the right hemisphere in Group 1 and over the left hemisphere in Group 2 (Figures 5D and 6D). The maximal current densities were located at 9, –72, 57 for Group 1 and –17, –86, 40 for Group 2.

### DISCUSSION

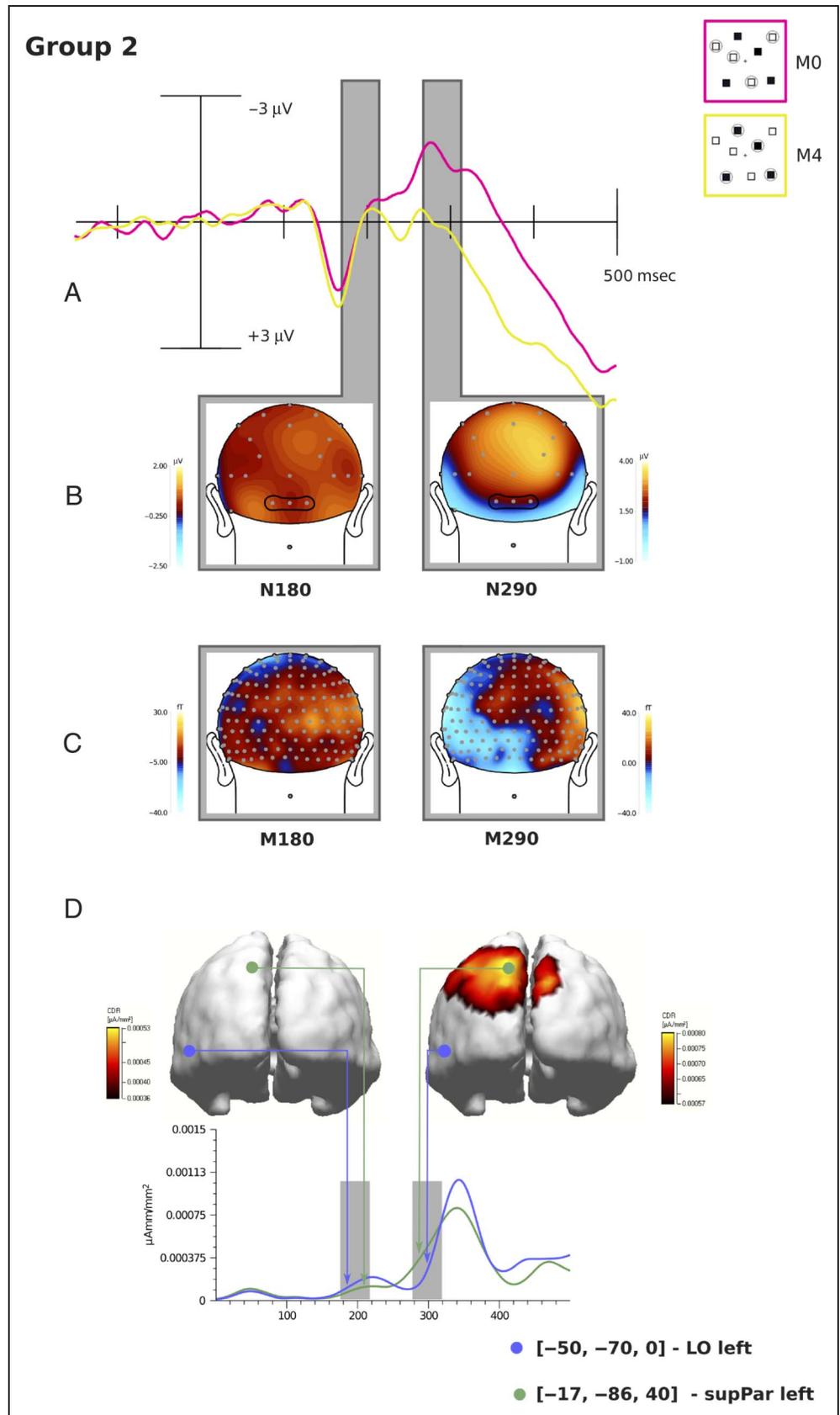
This study investigated visual processing mechanisms engaged during MOT by recording electrophysiological and neuromagnetic brain activity during task performance. The aim was to explore the hypothesis that MOT might be facilitated when participants perceive the global configuration or shape defined by the positions of the tracked objects, as proposed by Yantis (1992). In an experimental design that was free from sensory confounds, participants tracked four of eight moving items and were required to respond to a subsequent probe stimulus (luminance increment of four items) by pressing one of two buttons to indicate whether or not the probe matched all four of the tracked target items. Importantly, two distinct response patterns emerged from the analysis of the behavioral data. Whereas all participants responded quickly and accurately when the probe did not match any of the tracked items (match-0 condition), half of the participants were also fast and accurate when the probe matched all four of the tracked items (match-4 or full-match condition). These latter participants were assigned to Group 1, whereas participants who did not show such preferential responses to the full-match probes were assigned to Group 2. The rapid and accurate responses given by Group 1 participants to the full-match probes relative to the partial match probes suggests that they were using a perceptual strategy of monitoring the global shape configuration of the tracked target items. An alternative explanation could be the existence of a response bias given that 80% of the trials were not full-match trials. However, several studies (Klingstone & Klein, 1991; Lambert & Hockey, 1986; Laming, 1968) showed that a typical response bias would lead to slower and more erroneous responses to the unlikely stimulus. Here we observed the opposite pattern, which argues strongly against this explanation.

The ERP results provided converging evidence that Group 1 participants were monitoring the four tracked items as a global configuration. The full-match probes elicited an enlarged negative ERP at 170–190 msec (the N180 component) relative to the partial-match or no-match probes. This effect was absent in the Group 2 participants who did not preferentially respond to the full-match probes. Source localization of the enhanced N180 using

**Figure 5.** Estimated sources of N180/M180 and N290/M290 modulations in Group 1. Sources were estimated from the difference waveforms between match-4 and match-0 trials. (A) ERP waveforms on match-4 and match-0 trials. (B) Voltage topographies of the N180 and N290 difference waveforms. (C) ERF topographies of the M180 and M290 difference waveforms. (D) Estimated source of current densities based on concurrent fitting to ERP and ERF topographies for the N180/M180 and N290/M290 modulations.



**Figure 6.** Same as Figure 5 for Group 2.



combined ERP and neuromagnetic (ERF) field topographies estimated its neural generators to reside in LO, which has been implicated as a key area for perceptual encoding of shapes and objects (Kourtzi & Kanwisher, 2000; Grill-Spector et al., 1998; Malach et al., 1995). The behavioral and ERP evidence thus support the hypothesis that Group 1 participants were tracking the target items by monitoring the global shape configuration formed by their positions, and this tracking was accomplished via representations in LO. According to this hypothesis, the encoded representation of the tracked items' configuration in LO would facilitate the neural response to the probe stimulus that precisely matched that shape configuration and would allow for a rapid behavioral response to the full-match probe.

A subsequent ERP component elicited by the probe at 270–310 msec (the N290 component) showed a monotonic increase in amplitude as a function of decreasing probe–target congruity in both groups of participants. The sources of the N290 were estimated to lie in dorsal-medial parietal cortex (BA 7). The N290 amplitude appears to reflect the graded degree of mismatch between the locations of the tracked target items and the probe items. A similar late negative component (N270) has been described in situations where a relevant stimulus mismatches its perceptual representation in visual STM (VSTM; Astle, Scerif, Kuo, & Nobre, 2009; Wang, Wang, Cui, Tian, & Zhang, 2002). In the current MOT task, the N290 modulations suggest that location information for each of the individual tracked items was being stored in VSTM, and the N290 amplitude reflects the number of probe items that mismatch the VSTM representation. The behavioral and ERP results thus suggest that the MOT task can be carried out by two separate perceptual-cognitive strategies, one that involves monitoring of the global configuration or shape of the tracked items (reflected in the N180 in Group 1) and a second that involves monitoring the spatio-temporal properties of the individual target items (reflected in the N290 in both groups).

### **N180: Rapid Global Shape Analysis**

The enhanced N180 amplitude elicited by the full-match probes in Group 1 participants was localized to bilateral neural generators in LO. It is well documented that the LO region is critically involved in the extraction of shape information required for the perception of objects (Kourtzi & Kanwisher, 2000; Grill-Spector et al., 1998; Malach et al., 1995). Typically LO is activated by global shapes that are defined by some stimulus feature (Vinberg & Grill-Spector, 2008; Kourtzi & Kanwisher, 2000). The shape-defining feature dimension can be color, motion (Jiang et al., 2008; Schoenfeld, Woldorff, et al., 2003; Grill-Spector et al., 1998), or even illusory contours (Mendola, Dale, Fischl, Liu, & Tootell, 1999). Studies using ERPs have shown an enhanced negative component at 150–200 msec similar to the N180 reported here during the perception of a shape defined by illusory contours (Doniger et al., 2001;

Herrmann & Bosch, 2001). Moreover, this negative component elicited by illusory contours was shown to originate in the LO complex (Shpaner, Murray, & Foxe, 2009; Murray et al., 2002). These findings are in accord with the present results and with the proposition that this early LO modulation reflects a rapid analysis of the configurational properties of a stimulus array or object (Doniger et al., 2001).

Indeed, the tracking of the global configuration of multiple targets would seem to involve perceptual processes similar to those that segment shapes out of illusory contours. When the probe matches all four of the tracked items, it can facilitate the shape representation constructed from the imaginary contours that connect the positions of the tracked items. This is associated with a stronger neural response in LO reflected in the enhanced amplitude of the N180 component. Importantly, this effect was only evident for the full-match probes and only in the participants who also showed a corresponding behavioral advantage. These results are in line with observations that receptive fields in LO are relatively large and that LO activation can be highly selective to specific objects (Murray & Wojciulik, 2004; Grill-Spector et al., 1998). Accordingly, the altered shape associated with a partially matching probe would not engage the object representation built up for the tracked items and would not elicit an enhanced N180 component.

### **N290: Encoding of Individual Items**

The subsequent negativity peaking around 290 msec showed a gradual decrease in amplitude with increasing correspondence between target and probe in both groups. Consequently, this mismatch effect was independent of the presence of an amplitude effect in the preceding time range of the N180 modulation for the full match condition. The sources of this effect were localized to medial parietal cortex. Previous studies have reported negative amplitude modulations as a function of the degree of deviance between succeeding auditory stimuli (Breton, Ritter, Simson, & Vaughan, 1988; Sams, Alho, & Näätänen, 1983). This MMN has been interpreted to reflect an automatic change detection process of sequential presentations (Winkler, Schroger, & Cowan, 2001; Schroger, 1997) and could also be observed in visual tasks (Pazo-Alvarez, Cadaveira, & Amenedo, 2003). The amplitude of the visual MMN also increased with the amount of discrepancy between a current visual stimulus and a template held in memory (Stefanics, Kimura, & Czigler, 2011; Czigler, Balazs, & Winkler, 2002; Woods, Alho, & Algazi, 1992). In the same vein, Wang et al. (2002) reported a negativity in approximately the same time range (N270 effect) that reflected the perceptual difference between a present VSTM representation and an upcoming stimulus (Zhang, Wang, Li, Wang, & Tian, 2005; Zhang et al., 2001; Cui, Wang, Wang, Tian, & Kong, 2000). Hemodynamic activations in the intraparietal sulcus that correlated with the demands of MOT tasks were reported by several studies (Howe,

Horowitz, Morocz, Wolfe, & Livingstone, 2009; Xu & Chun, 2006; Culham, Cavanagh, & Kanwisher, 2001; Jovicich et al., 2001). This is well in line with the localization of the sources in the parietal cortex of the current study. Given its retinotopic organization (Walther & Koch, 2006; Silver, Ress, & Heeger, 2005), this region is ideally suited for the maintenance of a VSTM template with the spatio-temporal representation of stimuli. Importantly, locational information about attended stimuli was suggested to enter automatically into STM (Serenó & Amador, 2006; Tsal & Lavie, 1993) even when the information is irrelevant for the task (Serenó & Amador, 2006). Our results support this idea, as the parietal mismatch effect in the N290 time range was present in both groups. Whereas Group 2 could have used the processing underlying this effect for the selection of the response, the information would be redundant for Group 1 who could have selected their response earlier based on the processes underlying the N180 modulation (enhanced amplitude in the case of the full match).

### Impact on MOT

In general, the global aspects of a visual stimulus are processed preferential to local stimulus features (Navon, 1977). This might also apply to MOT. In the current experiment, about half of the participants (Group 1) exhibited a better performance in the MOT compared with the other half (Group 2). In the electrophysiological data, we observed two effects with a very different behavior. The late effect on the N290 component consisted of amplitude modulations as a factor of the number of probes matching with the tracked targets (mismatch effect). The early effect on the N180 component exhibited an amplitude modulation exclusively when all probe items matched the tracked targets (match effect). The N290 effect appears to reflect working memory processes, which is further supported by the presence of the effect in both groups and by the modeled sources localized in the parietal cortex. This effect supports the idea that for both participant groups the tracked objects are the units of attentional selection in the current MOT task (Scholl et al., 2001). The N180 effect, however, rather reflects processes related to the fact that the four items probe fall into a shape representation that was previously formed during tracking. The modeled source in the region LO for this match effect as well as the absence of the effect in the second group provides further support for this interpretation. This indicates that information about the overall spatial configuration of the items can be employed to facilitate their tracking and provide strong evidence at neurophysiological level.

### Acknowledgments

This work was supported by Deutsche Forschungsgemeinschaft Grant SCHO 1217/2-2, SFB779/TPA1, the U. S. National Science Foundation (BCS-1029084), National Institute of Mental Health

(1P50MH86385), and the Office of Naval Research (N 00014-07-1-0937).

Reprint requests should be sent to Christian Merkel, Department of Neurology, Otto-von-Guericke University, Leipziger Str. 44, 39120 Magdeburg, Germany, or via e-mail: christian.merkel@med.ovgu.de.

### 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, 1–10.
- Astle, D. E., Scerif, G., Kuo, B. C., & Nobre, A. C. (2009). Spatial selection of features within perceived and remembered objects. *Frontiers in Human Neuroscience*, 3, 6.
- Boehler, C. N., Tsotsos, J. K., Schoenfeld, M. A., Heinze, H. J., & Hopf, J. M. (2009). The center-surround profile of the focus of attention arises from recurrent processing in visual cortex. *Cerebral Cortex*, 19, 982–991.
- Bondarenko, R., Boehler, C. N., Stoppel, C. M., Heinze, H. J., Schoenfeld, M. A., & Hopf, J. M. (2012). Separable mechanisms underlying global feature-based attention. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 32, 15284–15295.
- Breton, F., Ritter, W., Simson, R., & Vaughan, H. G., Jr. (1988). The N2 component elicited by stimulus matches and multiple targets. *Biological Psychology*, 27, 23–44.
- Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, 9, 349–354.
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, 7, 249–253.
- Cui, L., Wang, Y., Wang, H., Tian, S., & Kong, J. (2000). Human brain sub-systems for discrimination of visual shapes. *NeuroReport*, 11, 2415–2418.
- Culham, J. C., Cavanagh, P., & Kanwisher, N. G. (2001). Attention response functions: Characterizing brain areas using fMRI activation during parametric variations of attentional load. *Neuron*, 32, 737–745.
- Czigler, I., Balazs, L., & Winkler, I. (2002). Memory-based detection of task-irrelevant visual changes. *Psychophysiology*, 39, 869–873.
- Doniger, G. M., Foxe, J. J., Schroeder, C. E., Murray, M. M., Higgins, B. A., & Javitt, D. C. (2001). Visual perceptual learning in human object recognition areas: A repetition priming study using high-density electrical mapping. *NeuroImage*, 13, 305–313.
- Drew, T., Horowitz, T. S., & Vogel, E. K. (2013). Swapping or dropping? Electrophysiological measures of difficulty during multiple object tracking. *Cognition*, 126, 213–223.
- Drew, T., Horowitz, T. S., Wolfe, J. M., & Vogel, E. K. (2012). Neural measures of dynamic changes in attentive tracking load. *Journal of Cognitive Neuroscience*, 24, 440–450.
- Drew, T., McCollough, A. W., Horowitz, T. S., & Vogel, E. K. (2009). Attentional enhancement during multiple-object tracking. *Psychonomic Bulletin & Review*, 16, 411–417.
- Driver, J., & Baylis, G. C. (1989). Movement and visual attention: The spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 448–456.
- Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition*, 80, 61–95.
- Dubois, J., Hamker, F. H., & VanRullen, R. (2009). Attentional selection of noncontiguous locations: The spotlight is only transiently “split”. *Journal of Vision*, 9, 1–11.

- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*, 501–517.
- Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 32–39.
- 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.
- 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.
- Fuchs, M., Wagner, M., Wischmann, H. A., Kohler, T., Theissen, A., Drenckhahn, R., et al. (1998). Improving source reconstructions by combining bioelectric and biomagnetic data. *Electroencephalography and Clinical Neurophysiology*, *107*, 93–111.
- Grill-Spector, K., Kushnir, T., Edelman, S., Itzhak, Y., & Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron*, *21*, 191–202.
- Herrmann, C. S., & Bosch, V. (2001). Gestalt perception modulates early visual processing. *NeuroReport*, *12*, 901–904.
- Hillyard, S. A., Teder-Salejari, W. A., & Munte, T. F. (1998). Temporal dynamics of early perceptual processing. *Current Opinion in Neurobiology*, *8*, 202–210.
- Hopf, J. M., Boehler, C. N., Luck, S. J., Tsotsos, J. K., Heinze, H. J., & Schoenfeld, M. A. (2006). Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings of the National Academy of Sciences, U.S.A.*, *103*, 1053–1058.
- Howe, P. D., Horowitz, T. S., Morocz, I. A., Wolfe, J., & Livingstone, M. S. (2009). Using fMRI to distinguish components of the multiple object tracking task. *Journal of Vision*, *9*, 1–11.
- Jiang, Y., Boehler, C. N., Nonnig, N., Duzel, E., Hopf, J. M., Heinze, H. J., et al. (2008). Binding 3-D object perception in the human visual cortex. *Journal of Cognitive Neuroscience*, *20*, 553–562.
- Jovicich, J., Peters, R. J., Koch, C., Braun, J., Chang, L., & Ernst, T. (2001). Brain areas specific for attentional load in a motion-tracking task. *Journal of Cognitive Neuroscience*, *13*, 1048–1058.
- Klingstone, A., & Klein, R. (1991). Combining shape and position expectancies: Hierarchical processing and selective inhibition. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 512–519.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, *20*, 3310–3318.
- Lambert, A., & Hockey, R. (1986). Selective attention and performance with a multidimensional visual display. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 484–495.
- Laming, D. R. J. (1968). *Information theory of choice-reaction times*. London: Academic Press.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*, 19–40.
- Makovski, T., Vazquez, G. A., & Jiang, Y. V. (2008). Visual learning in multiple-object tracking. *PLoS One*, *3*, e2228.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *92*, 8135–8139.
- Mangun, G. R., & Hillyard, S. A. (1988). Spatial gradients of visual attention: Behavioral and electrophysiological evidence. *Electroencephalography and Clinical Neurophysiology*, *70*, 417–428.
- Martinez, A., Ramanathan, D. S., Foxe, J. J., Javitt, D. C., & Hillyard, S. A. (2007). The role of spatial attention in the selection of real and illusory objects. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*, 7963–7973.
- McMains, S. A., & Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron*, *42*, 677–686.
- Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality*, *60*, 117–174.
- Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, *19*, 8560–8572.
- Muller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature*, *424*, 309–312.
- Murray, M. M., Wylie, G. R., Higgins, B. A., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). The spatiotemporal dynamics of illusory contour processing: Combined high-density electrical mapping, source analysis, and functional magnetic resonance imaging. *Journal of Neuroscience*, *22*, 5055–5073.
- Murray, S. O., & Wojciulik, E. (2004). Attention increases neural selectivity in the human lateral occipital complex. *Nature Neuroscience*, *7*, 70–74.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383.
- Niebergall, R., Khayat, P. S., Treue, S., & Martinez-Trujillo, J. C. (2011). Multifocal attention filters targets from distracters within and beyond primate MT neurons' receptive field boundaries. *Neuron*, *72*, 1067–1079.
- O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature*, *401*, 584–587.
- Pazo-Alvarez, P., Cadaveira, F., & Amenedo, E. (2003). MMN in the visual modality: A review. *Biological Psychology*, *63*, 199–236.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25.
- Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*, 65–97.
- Pylyshyn, Z. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition*, *11*, 801–822.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 179–197.
- Rensink, R. A., & Enns, J. T. (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review*, *102*, 101–130.
- Sams, M., Alho, K., & Näätänen, R. (1983). Sequential effects on the ERP in discriminating two stimuli. *Biological Psychology*, *17*, 41–58.
- Schneider, W. X. (1993). Space-based visual attention models and object selection: Constraints, problems, and possible solutions. *Psychological Research*, *56*, 35–43.
- Schoenfeld, M. A., Hopf, J. M., Martinez, A., Mai, H. M., Sattler, C., Gasde, A., et al. (2007). Spatio-temporal analysis of feature-based attention. *Cerebral Cortex*, *17*, 2468–2477.
- Schoenfeld, M. A., Tempelmann, C., Martinez, A., Hopf, J. M., Sattler, C., Heinze, H. J., et al. (2003). Dynamics of feature binding during object-selective attention. *Proceedings of the National Academy of Sciences, U.S.A.*, *100*, 11806–11811.

- Schoenfeld, M. A., Woldorff, M., Duzel, E., Scheich, H., Heinze, H. J., & Mangun, G. R. (2003). Form-from-motion: MEG evidence for time course and processing sequence. *Journal of Cognitive Neuroscience*, *15*, 157–172.
- 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.
- Schroger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology*, *34*, 245–257.
- Sereno, A. B., & Amador, S. C. (2006). Attention and memory-related responses of neurons in the lateral intraparietal area during spatial and shape-delayed match-to-sample tasks. *Journal of Neurophysiology*, *95*, 1078–1098.
- Shpaner, M., Murray, M. M., & Foxe, J. J. (2009). Early processing in the human lateral occipital complex is highly responsive to illusory contours but not to salient regions. *European Journal of Neuroscience*, *30*, 2018–2028.
- Silver, M. A., Ress, D., & Heeger, D. J. (2005). Topographic maps of visual spatial attention in human parietal cortex. *Journal of Neurophysiology*, *94*, 1358–1371.
- Stefanics, G., Kimura, M., & Czigler, I. (2011). Visual mismatch negativity reveals automatic detection of sequential regularity violation. *Frontiers in Human Neuroscience*, *5*, 46.
- Tombu, M., & Seiffert, A. E. (2008). Attentional costs in multiple-object tracking. *Cognition*, *108*, 1–25.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*, 15–48.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*, 107–141.
- Tsal, Y., & Lavie, N. (1993). Location dominance in attending to color and shape. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 131–139.
- Vinberg, J., & Grill-Spector, K. (2008). Representation of shapes, edges, and surfaces across multiple cues in the human visual cortex. *Journal of Neurophysiology*, *99*, 1380–1393.
- Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks: The Official Journal of the International Neural Network Society*, *19*, 1395–1407.
- Wang, Y., Wang, H., Cui, L., Tian, S., & Zhang, Y. (2002). The N270 component of the event-related potential reflects supramodal conflict processing in humans. *Neuroscience Letters*, *332*, 25–28.
- Winkler, I., Schroger, E., & Cowan, N. (2001). The role of large-scale memory organization in the mismatch negativity event-related brain potential. *Journal of Cognitive Neuroscience*, *13*, 59–71.
- Woods, D. L., Alho, K., & Algazi, A. (1992). Intermodal selective attention. I. Effects on event-related potentials to lateralized auditory and visual stimuli. *Electroencephalography and Clinical Neurophysiology*, *82*, 341–355.
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*, 91–95.
- Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*, 295–340.
- Zhang, X., Wang, Y., Li, S., Wang, L., & Tian, S. (2005). Distinctive conflict processes associated with different stimulus presentation patterns: An event-related potential study. *Experimental Brain Research*, *162*, 503–508.
- Zhang, Y., Wang, Y., Wang, H., Cui, L., Tian, S., & Wang, D. (2001). Different processes are involved in human brain for shape and face comparisons. *Neuroscience Letters*, *303*, 157–160.