

# Preparatory Activity in Posterior Temporal Cortex Causally Contributes to Object Detection in Scenes

Reshanne R. Reeder<sup>1,2</sup>, Francesca Perini<sup>1</sup>, and Marius V. Peelen<sup>1</sup>

## Abstract

■ Theories of visual selective attention propose that top-down preparatory attention signals mediate the selection of task-relevant information in cluttered scenes. Neuroimaging and electrophysiology studies have provided correlative evidence for this hypothesis, finding increased activity in target-selective neural populations in visual cortex in the period between a search cue and target onset. In this study, we used online TMS to test whether preparatory neural activity in visual cortex is causally involved in naturalistic object detection. In two experiments, participants detected the presence of object categories (cars, people) in a diverse set of photographs of real-world scenes. TMS was applied over a region in posterior tem-

poral cortex identified by fMRI as carrying category-specific preparatory activity patterns. Results showed that TMS applied over posterior temporal cortex before scene onset (−200 and −100 msec) impaired the detection of object categories in subsequently presented scenes, relative to vertex and early visual cortex stimulation. This effect was specific to category level detection and was related to the type of attentional template participants adopted, with the strongest effects observed in participants adopting category level templates. These results provide evidence for a causal role of preparatory attention in mediating the detection of objects in cluttered daily-life environments. ■

## INTRODUCTION

Visual selective attention serves to prioritize the processing of objects in our environment that are relevant to current goals, such as when we look out for cars before crossing the street. Theories of visual attention propose that attentional selection in cluttered displays is guided by internal descriptions of task-relevant information, or “attentional templates” (Duncan & Humphreys, 1989; Wolfe, Cave, & Franzel, 1989), with the degree to which an object matches this internal description influencing the likelihood that the object is attended.

Although the concept of an attentional template is intuitively appealing, its specific neural implementation and functional significance remain unclear, particularly in naturalistic vision (Peelen & Kastner, 2014). One hypothesis is that preparing to detect a target stimulus results in the preactivation or priming of neurons that represent that stimulus in visual cortex (Desimone & Duncan, 1995). This preactivation would result in a competitive bias in favor of template-matching stimuli, perhaps by lowering the sensory input needed for these neurons to fire. Single-cell recording and human neuroimaging studies have provided evidence for this hypothesis by showing prestimulus increases in target-selective populations of neurons and voxels (Peelen & Kastner, 2011; Chelazzi, Miller, Duncan, & Desimone, 1993). It is

unclear, however, whether this activity is causally involved in the efficient selection of subsequently presented targets or whether it reflects processes that are unrelated to target detection. Indeed, an alternative (not mutually exclusive) possibility is that attentional templates are implemented in regions outside visual cortex, for example, in regions of pFC implicated in working memory (Bansal et al., 2014; Miller, Erickson, & Desimone, 1996), with attentional selection mediated by feedback from these regions only after the onset of a search display.

In this study, we used online TMS to test whether preparatory neural activity in high-level visual cortex is causally involved in the detection of targets in cluttered scenes. In two experiments, participants detected the presence of object categories (cars, people) in a diverse set of photographs of real-world scenes while TMS was applied before scene onset. The target region for TMS was derived from a recent fMRI study testing for preparatory attention signals using the same search task (Peelen & Kastner, 2011). This study showed that activity patterns in a region of right posterior temporal cortex (pTC) during the preparatory phase of naturalistic visual search carried information about the to-be-detected object category, resembling category level attentional templates. Importantly, the strength of this effect in pTC was strongly correlated with participants’ accuracy, suggesting that preparatory activity in this region was behaviorally relevant. By contrast, the category selectivity of preparatory activity patterns in

<sup>1</sup>University of Trento, <sup>2</sup>Otto-von-Guericke University, Magdeburg, Germany

early visual cortex (EVC) was negatively correlated with accuracy. These individual differences at the neural level were linked to the use of different types of templates (indexed by a questionnaire), with “good” searchers reporting to use more high-level templates (searching for high-level view-invariant features of categories, represented in pTC) and “poor” searchers reporting to use more low-level templates (searching for low-level features associated with the target categories).

These findings led to specific predictions for the current TMS study, tested in two experiments. Experiment 1 tested (1) whether TMS over pTC before scene onset (“prescene TMS”) negatively affects detection performance relative to prescene TMS over vertex and (2) whether this effect is stronger for participants adopting relatively high-level categorical attentional templates. Experiment 2 focused on the regional and task specificity of prescene TMS, testing whether prescene TMS over pTC, relative to prescene TMS over EVC, selectively affects detection at the category level relative to detection based on lower level features.

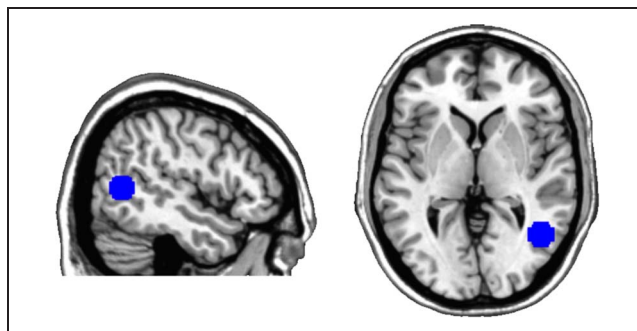
## METHODS

### Participants

Fifty-two healthy right-handed undergraduate and graduate students from the University of Trento participated in the study. Sixteen of these participated in Experiment 1 (eight women; aged 21–31 years, mean = 25.5 years), and 40 (four of them also participated in Experiment 1) participated in Experiment 2 (25 women; aged 19–34 years, mean = 24.6 years). The large sample size of Experiment 2 was related to the counterbalancing of condition order (see Experiment 2 Procedure). One participant of Experiment 1 was excluded because of low overall performance (mean accuracy was 2 SDs below the group mean). All participants completed a screening questionnaire to ensure that they met the safety criteria to undergo TMS experimentation. Participants received stimulation to the target areas before the task to expose them to the sensation of TMS and to ensure that stimulation was comfortable for all stimulation sites. All participants were comfortable with the TMS procedure and received monetary compensation for their participation. All participants provided written, informed consent before taking part in the experiment. The study was approved by the human research ethics committee of the University of Trento and adhered to the tenets of the Declaration of Helsinki.

### General TMS Methods

During the experiment, participants sat in a straight-backed chair and rested their heads on a chin rest to minimize body and head movement during experimentation. Once the stimulation site was located, the TMS coil was stabilized against the participant’s head using a metal



**Figure 1.** Region in pTC stimulated with TMS in Experiments 1 and 2. pTC location was based on fMRI results of Peelen and Kastner (2011).

arm. TMS pulses were delivered with a Magstim Rapid stimulator with a 75-mm MCF-B65 Butterfly coil (Experiment 1) and a Magstim Rapid stimulator with a 50-mm D70 Alpha coil (Experiment 2). The hand area of the motor cortex was first localized in the left hemisphere. The visual resting motor threshold (MT) was determined for each participant based on the minimum stimulation intensity needed to produce a visible right-hand twitch in at least 5 of 10 pulses. Stimulation intensity for experimentation was set at 120% of each participant’s MT. MT, rather than phosphene threshold, was used to determine stimulation intensity, as MT could be determined in all participants (unlike phosphene threshold; see EVC localization below).

For each participant, pTC was localized using anatomical brain scans acquired for previous (unrelated) fMRI experiments. Anatomical data were normalized to Talairach space to localize pTC based on the Talairach coordinates reported in Peelen and Kastner (2011; peak:  $x, y, z = 46, -58, 8$ ; see Figure 1 for an illustration of the pTC coordinates mapped on a sample brain). This region is located in pTC, about 2–3 cm anterior to retinotopically defined regions LO1 ( $y = -89$ ) and LO2 ( $y = -82$ ; Larsson & Heeger, 2006) and about 1 cm anterior to the peaks of object-selective LO ( $y = -71$ ), body-selective EBA ( $y = -67$ ), and motion-selective MT+ ( $y = -65$ ; Downing, Wiggett, & Peelen, 2007). Once pTC was located, brain images were transformed back to native space and used to position the coil during the experiment. pTC was located for each participant using Zebris neuronavigation software, and its position was marked on the scalp with a permanent marker. The coil was placed over pTC with the handle pointed toward the back of the head. The coil was turned slightly clockwise or counterclockwise to produce inferior-to-superior current flow to find the best stimulation angle that minimized muscle twitching and eye blinking.

In Experiment 1, effects of pTC stimulation were compared with effects of vertex stimulation. Vertex was located by placing the coil centered between the two cerebral hemispheres on the top of the scalp halfway between the inion and the nasion. Single-pulse TMS was

administered at one of five time points relative to scene onset (−200, −100, 0, 100, and 200 msec). Stimulation times were randomized with an equal number of trials for each of the stimulation times within a block. In Experiment 2, effects of pTC stimulation were compared with effects of EVC stimulation. EVC was located by placing the coil 2 cm above theinion pointed inferiorly and adjusting it slightly to the left and right until stimulation evoked the perception of phosphenes (flashes of light) in blindfolded participants. Stimulation intensity was varied until participants reported seeing phosphenes. For participants who did not report seeing phosphenes (15/40 participants), EVC was located by placing the coil 2 cm above theinion pointed inferiorly. Double-pulse TMS was administered starting 200 msec before scene onset (i.e., pulses were delivered at −200 and −100 msec relative to scene onset) on half of the trials (randomly selected). No TMS was delivered on the other half of the trials to prevent the coil from overheating during an experimental session. No significant accuracy differences between conditions were found on trials without TMS ( $p > .18$ , for all tests). Overall performance was faster for TMS (703 msec) than for no-TMS trials (736 msec;  $t(39) = 6.79, p < .001$ , two tailed), likely reflecting an alerting or arousal effect because of the clicking sound that accompanies the prescene TMS pulses. Perhaps, because of this alerting effect, accuracy performance was also slightly, although reliably, higher for TMS (86.1%) than no-TMS (85.2%) trials ( $t(39) = 2.19, p = .035$ , two tailed). This artifact renders the no-TMS trials unsuitable as a control condition (Duecker & Sack, 2013), and all further analyses focused on comparing performance between TMS conditions.

## Stimulus Presentation

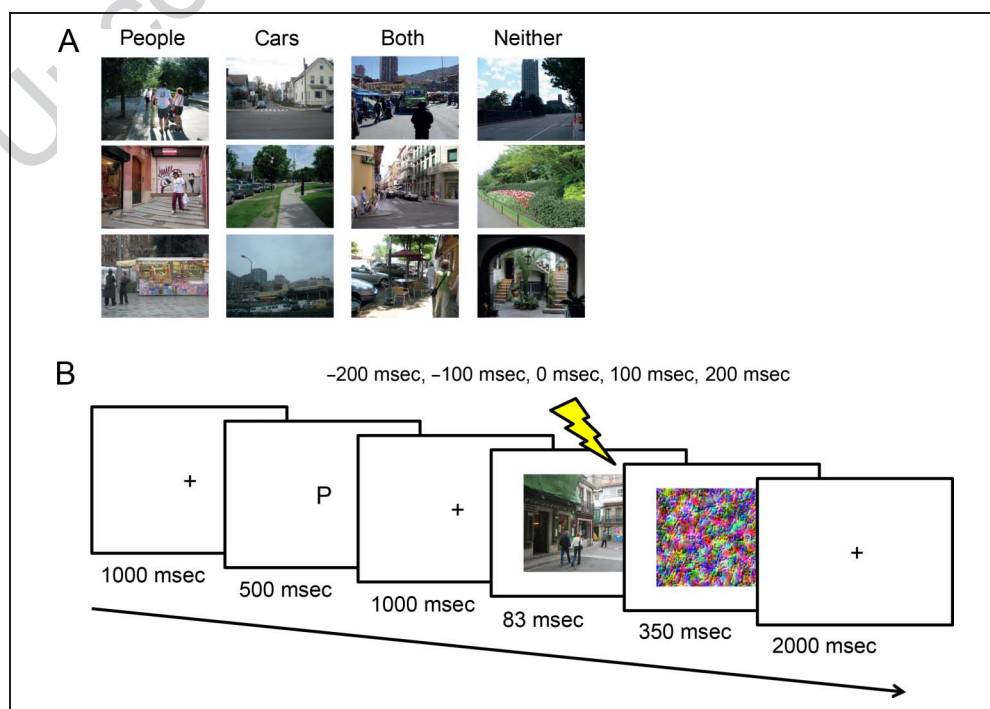
Stimuli were presented on a 17-in. gamma-corrected DELL 1908FP-BLK monitor with a screen resolution of  $1280 \times 960$  pixels and a 60-Hz refresh rate (Experiment 1) and a 22-in. Dell E228 WFP monitor with a screen resolution of  $1680 \times 1050$  pixels and a 60-Hz refresh rate (Experiment 2). In both experiments, stimuli were presented using A Simple Framework (Schwarzbach, 2011), based on the Psychophysics Toolbox for MATLAB.

## Experiment 1 Procedure

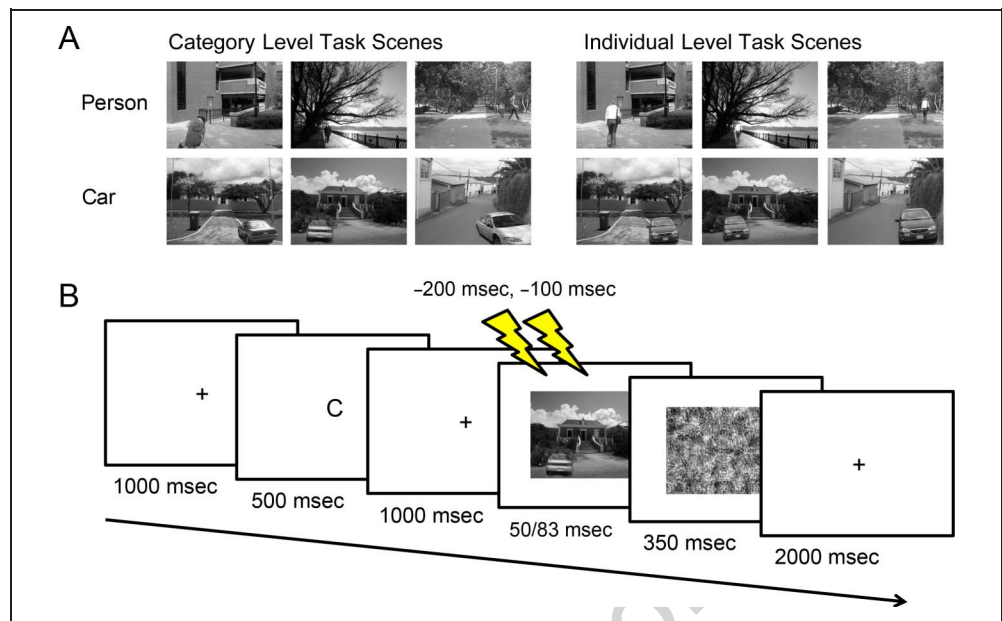
Participants were cued to detect cars and people in diverse photographs containing cars, people, both cars and people, or other objects in a natural scene context (see Figure 2A for examples). Stimuli were 640 color photographs of real-world scenes, obtained from an online database and used in previous fMRI experiments (Peelen & Kastner, 2011; Peelen, Fei-Fei, & Kastner, 2009). Photographs could contain people (160), cars (160), both people and cars (160), or no people or cars (160).

Figure 2B provides a schematic of the experimental paradigm. A trial began with the presentation of a centrally presented fixation cross (visual angle =  $0.81^\circ$ ) for 1000 msec, followed by a centrally presented single letter cue (visual angle =  $1.83^\circ$ ) for 500 msec, indicating the target category for that trial: “P” for “persona” or “M” for “macchina” (the Italian words for person and car, respectively). After the cue, another fixation cross appeared for 1 sec, followed by a centrally presented photograph of a real-world scene (visual angle =  $16.08^\circ \times 12.18^\circ$ ) for 83 msec. All four scene types were presented an equal

**Figure 2.** Overview of experimental procedure of Experiment 1. (A) Examples of scene pictures presented. Scenes could show people, cars, both people and cars, or neither people nor cars. (B) Schematic of trial structure (not drawn to scale). Participants were cued to detect people or cars in briefly presented scenes. A letter cue indicated the target category for each trial. Single-pulse TMS was applied over pTC or vertex at one of the five time points relative to scene onset.



**Figure 3.** Overview of experimental procedure of Experiment 2. (A) Examples of scene pictures. In the category level task, different car and person exemplars were presented in each trial, whereas in the individual level task, the car and person exemplars were held constant across trials. (B) Schematic of trial structure (not drawn to scale). Participants were cued to detect people or cars in briefly presented scenes. A letter cue indicated the target category for each trial. To equate performance, presentation time was 50 msec for the individual level task and 83 msec for the category level task. Double-pulse TMS was applied over pTC or EVC before scene onset (−200 and −100 msec).



number of times within a block (20 each), with presentations of the “P” and “M” cues distributed evenly across the scene types. Scenes were backward masked for 350 msec, after which a final fixation cross appeared for 2 sec, for a total trial duration of approximately 4.9 sec. Participants were instructed to respond whether the cued object (person or car) appeared in the scene by pressing 1 for “yes” or 2 for “no” on the keyboard number pad. All scenes were unique, such that participants never saw the same scene twice. All participants completed a practice block followed by eight experimental blocks of 80 trials each.

TMS was applied over pTC or vertex alternating every two blocks (e.g., two blocks of pTC stimulation, two blocks of vertex stimulation, two blocks of pTC). The order in which pTC and vertex were stimulated was counter-balanced across participants.

After the experiment, participants filled out the attentional template questionnaire introduced by Peelen and Kastner (2011). The questionnaire consists of 10 statements that probe the manner in which participants used the cue to prepare for object detection—whether they used relatively high-level attentional templates (consisting of general and view-invariant features of the target category) or relatively low-level attentional templates (consisting of specific features or exemplars of the target category). Participants rated their agreement with each statement on a 5-point scale, with 1 indicating that the participant fully disagreed and 5 indicating that the participant fully agreed with the statement. Examples of high-level template statements are “After the car cue I anticipated detecting cars seen from multiple angles rather than from one angle” and “After the person cue I formed a general idea of what a person in the scene may look like.” Examples of low-level template statements are “After the car cue I looked out for horizontal things that

were about the size of a car” and “After the person cue I imagined persons with a prototypical posture as seen from the front.” The full list of the statements can be found in Peelen and Kastner (2011). For each participant, the mean rating of the low-level template statements was subtracted from the mean rating of the high-level template statements, such that more positive scores indicate the use of relatively high-level attentional templates (Peelen & Kastner, 2011).

### Experiment 2 Procedure

Participants were cued to detect the presence of cars and people in real-world scenes. There were two scene sets, presented in different blocks. In the category level condition, the target object was different in each trial, such that participants had to look for category diagnostic features that were common to the exemplars to perform the task. By contrast, in the individual level condition, the target object was the same across all trials, such that participants could look for exemplar-specific features to perform the task (see Figure 3A). Stimuli were photographs of real-world scenes, obtained from an online database (Russell, Torralba, Murphy, & Freeman, 2008), converted to gray-scale. Target objects (cars, people) for the category level condition consisted of 48 different images of people and 48 different images of cars that were manually inserted into 96 scenes (Figure 3A); no people or cars were inserted in the remaining 96 scenes; 48 of these were presented in the category level condition, and 48 were presented in the individual level condition. For the individual level condition, person and car images were inserted into the 96 scenes that were also used in the category level condition. Importantly, in the individual level condition, the inserted person and car images were of the same person and car,



repeated in 48 scenes each (Figure 3A). Different person and car exemplars were used across participants, but the same exemplar was presented within a participant. Person and car images were placed in natural locations within a scene (i.e., a person could appear on a staircase, and a car could appear in a driveway). Targets were placed in various parts of the scene (far or near, to the left or right) and thus could appear large or small depending on the appropriateness of the scale. Category level and individual level targets were matched by location and size (i.e., the person in Scene 1 of the category level condition would be in the same location and of the same size as the person in Scene 1 of the individual level condition; Figure 3A).

Figure 3B shows a schematic of the experimental paradigm. A trial began with the presentation of a centrally presented fixation cross (visual angle =  $0.81^\circ$ ) for 1000 msec, followed by a centrally presented single letter cue (visual angle =  $1.83^\circ$ ) for 500 msec, indicating the target category for that trial: “P” for “persona” or “M” for “macchina” (the Italian words for person and car, respectively). After the cue, another fixation cross appeared for 1 sec, followed by a centrally presented photograph of a real-world scene (visual angle =  $16.08^\circ \times 12.18^\circ$ ) for 83 msec in the category level blocks or 50 msec in the individual level blocks; these presentation times were determined after extensive behavioral piloting (with different participants from those tested in the TMS experiment) with the aim to match the difficulty of the two blocks. There was an equal number (16) of scenes containing a person, a car, or neither person nor car within a block, and the two letter cues (“P” and “M”) appeared an equal number of times in each of the different scene types. Scenes were backward masked for 350 msec, after which a final fixation cross appeared for 2 sec, for a total trial duration of approximately 4.9 sec. Participants were instructed to respond whether the cued object (person or car) appeared in the scene by pressing 1 for “yes” or 2 for “no” on the keyboard number pad as fast and accurately as possible. In the category level blocks, participants never saw the same person or car twice within a block, whereas in the individual level blocks, the same person and car were repeated throughout a block.

All participants completed one experimental session in which they were given a practice block followed by 12 experimental blocks of 48 trials each. Participants performed six consecutive category level blocks and six consecutive individual level blocks, with three consecutive blocks of pTC stimulation and three consecutive blocks of EVC stimulation for each task. The same scenes were presented for each of the two stimulation sites. The order of the stimulated regions (EVC–pTC or pTC–EVC) was the same for category level and individual level conditions for a single participant, and this order was counterbalanced across participants. Condition order was also counterbalanced across participants for 10 participants in each region/condition order (i.e., if 1 = pTC stimulation in the category level condition, 2 = EVC stimulation

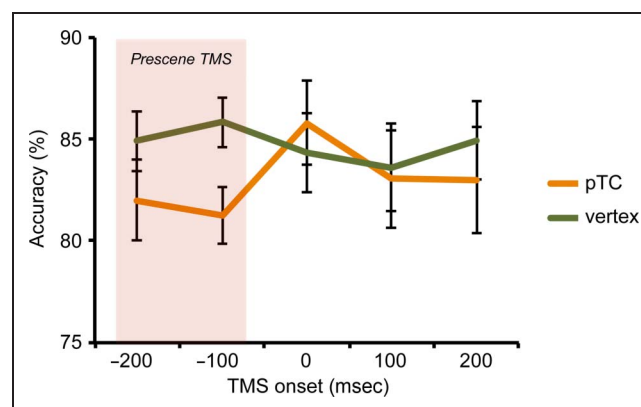
in the category level condition, 3 = pTC stimulation in the individual level condition, and 4 = EVC stimulation in the individual level condition, then there were 10 participants each in the following task orders: 1-2-3-4, 2-1-4-3, 3-4-1-2, and 4-3-2-1). Because the first 20 participants were tested with two of the four possible condition orders, we tested another 20 participants to complete the full counterbalancing.

## RESULTS

### Experiment 1

Participants were cued to detect people and cars in briefly presented photographs of real-world scenes while single-pulse TMS was delivered over pTC or vertex at different time points relative to scene onset (Figure 2B; see Methods). After the experiment, participants filled out a questionnaire that probed how they had used the cues to prepare for target detection—whether they had used relatively low- or high-level attentional templates (see Methods). The critical trials were those in which TMS was applied before scene onset (–200 and –100 msec), testing the causal role of preparatory activity in object detection.

A significant effect of TMS over pTC, relative to vertex, was found when TMS was applied 100 msec before scene onset ( $t(14) = -2.69, p = .009$ , one tailed; Figure 4). The corresponding effect at –200 msec approached significance ( $t(14) = -1.64, p = .06$ , one tailed). There were no significant differences between pTC and vertex when TMS was applied at or after scene onset (0, 100, and 200 msec;  $p > .3$ , for all tests). Prescene TMS over pTC, averaged across the –200 and –100 msec stimulation onsets (shaded area in Figure 4), significantly reduced accuracy relative to prescene TMS over vertex ( $t(14) = -3.03, p = .0045$ , one tailed).



**Figure 4.** Results of Experiment 1, showing effects of TMS at different time points (relative to scene onset) on category detection performance in briefly presented real-world scenes. Prescene TMS over pTC, relative to vertex, significantly reduced detection accuracy. Error bars indicate SEM.

Analysis of RT revealed no significant differences between pTC and vertex stimulation for any of the onset conditions ( $p > .26$ , for all tests).

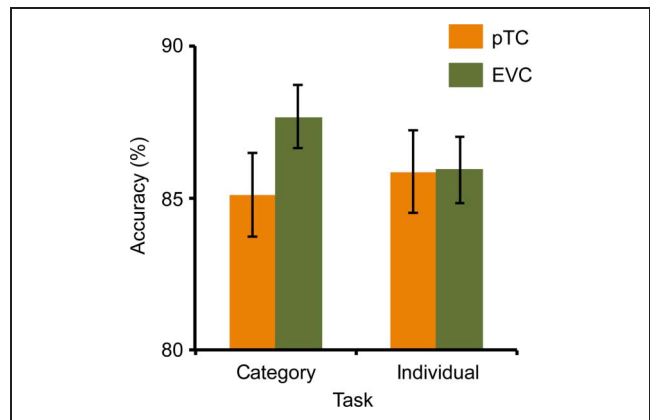
Next, we turned to the questionnaire data to test our second hypothesis: that the effect of prescene TMS (averaged across the  $-200$ - and  $-100$ -msec stimulation conditions; shaded area in Figure 4) over pTC is strongest for participants who adopted relatively high-level attentional templates (e.g., looking for cars at multiple angles and locations). Confirming this hypothesis, the effect of prescene TMS over pTC (relative to prescene TMS over vertex) was negatively correlated with template level ( $r(13) = -.576, p = .012$ , one tailed), reflecting stronger effects of pTC-TMS in participants who reported to use relatively high-level templates. To ensure that this correlation was not mediated by performance differences, we performed a partial correlation analysis between strategy and TMS effect, controlling for overall performance in the absence of TMS (average of vertex conditions). The correlation between strategy and TMS effect remained significant ( $r(13) = -.632, p = .006$ , one tailed).

Analysis of RT data showed no significant relationship between the effect of prescene TMS over pTC (relative to vertex) and template level ( $r(13) = .161, p = .58$ , two tailed).

## Experiment 2

Experiment 2 was conducted to test the hypothesis that prescene TMS over pTC, relative to prescene TMS over EVC, selectively affects detection at the category level relative to detection based on lower level features. To test this, participants performed a cued category detection task similar to that of Experiment 1. Importantly, however, in this experiment, we manipulated the features that were informative for the task. In the category level condition, scenes contained different exemplars of the target category on every trial, requiring category level attentional templates. By contrast, in the individual level condition, the same person or car was presented throughout a block, allowing for detection based on lower level features. In behavioral pilots, we found that the individual level condition was indeed easier than the category level condition, indicating that participants made use of the predictable features to optimize their performance. On the basis of these pilot data, we adjusted the presentation times to equate the difficulty of the two tasks. The two tasks were matched by presenting scenes for 83 msec in the category level task and for 50 msec in the individual level task (category level task: 86.4%, individual level task: 85.9%;  $F(1, 39) = 0.34, p = .56$ , two tailed). Double-pulse TMS was delivered before scene onset ( $-200$  and  $-100$  msec) to disrupt preparatory processing in pTC or EVC, in different blocks.

Confirming our hypothesis and replicating Experiment 1, prescene TMS over pTC reduced accuracy on the category level detection task, relative to prescene TMS over



**Figure 5.** Results of Experiment 2, showing the effect of prescene TMS (double pulse at  $-200$  and  $-100$  msec) on performance in category level and individual level detection tasks. TMS over pTC, relative to TMS over EVC, significantly reduced detection accuracy in the category level task.

EVC ( $t(39) = -2.34, p = .012$ , one tailed; Figure 5). By contrast, for the individual level task, prescene TMS over pTC had no effect ( $t(39) = -0.05, p = .96$ ), resulting in a significant Task (category level, individual level)  $\times$  Region (pTC, EVC) interaction ( $F(1, 39) = 3.2, p = .041$ , one tailed).

Analysis of RT data confirmed these results, showing slower responses in the category level task when TMS was applied over pTC relative to EVC ( $t(39) = 2.93, p = .003$ , one tailed). TMS had the opposite effect in the individual level task, with TMS over pTC speeding up responses relative to TMS over EVC ( $t(39) = -2.08, p = .022$ , one tailed); the Task (category, individual)  $\times$  Region (pTC, EVC) interaction was significant ( $F(1, 39) = 14.6, p < .001$ , one tailed).

## DISCUSSION

Theoretical and empirical studies have suggested an important role for top-down attentional templates in guiding real-world attentional selection (for reviews, see Peelen & Kastner, 2014; Wolfe, Vo, Evans, & Greene, 2011), but the neural implementation and functional significance of such templates have remained elusive. In this study, we used online TMS to show that preparatory attentional templates in high-level visual cortex causally contribute to the efficient detection of objects in natural scenes. In both experiments, TMS delivered over pTC before scene onset disrupted the detection of object categories in subsequently presented scenes, indicating that preparatory processing in this region supported the selection of task-relevant categorical information. Moreover, the effect of prescene TMS was related to the type of template participants adopted, with TMS over pTC specifically interfering with higher level categorical templates (Experiment 1) and in category level detection tasks (Experiment 2).

The current experiments were motivated by the findings of a previous study that used fMRI to investigate preparatory attention during the same naturalistic search task (Peelen & Kastner, 2011). This fMRI study revealed a pTC region in which the category specificity of attentional templates was strongly correlated with behavioral performance. The current results show that interfering with preparatory activity in this region by means of TMS disrupts category level detection performance. Interestingly, as in the previous fMRI study, this effect was most strongly observed in participants reporting to use category level templates. Using a task manipulation, Experiment 2 provides further evidence for a specific role of pTC in category level search. The finding that pTC represents relatively high-level object representations is in accordance with its position in the visual system hierarchy. The findings of a correlation between template level and the TMS effect (Experiment 1) and the specificity of TMS to category level detection tasks (Experiment 2) are important in that they rule out the possibility that the performance impairments after TMS over pTC were because of TMS-related artifacts (e.g., eye blinks or muscle twitches).

In Experiment 1, TMS over pTC after scene onset did not significantly interfere with performance. This could indicate a dissociation between the maintenance of a top-down attentional set (critically involving pTC) and the visual analysis of the scene itself (not critically involving pTC). However, there may be several other reasons for the absence of a postscene TMS effect. Perhaps, most importantly, the specific times at which TMS was delivered relative to scene onset in the current study (0, 100, and 200 msec) may have missed a critical time window at which this region is involved in visual processing. For example, a recent study investigating animal detection in scenes found that TMS over lateral occipital cortex affected subjective perception when applied 150 msec after scene onset, but not when applied 90 or 210 msec after scene onset (Koivisto, Railo, Revonsuo, Vanni, & Salminen-Vaparanta, 2011). In contrast, the critical time window of prescene TMS for disrupting preparatory templates would be expected to be broader than this. Another possibility for the lack of a significant postscene TMS effect is that the effect of stimulation to pTC was simply too weak to disrupt visually driven processing, with internally generated preparatory activity being more susceptible to interference. More work is needed to clarify the association or dissociation between template-related and stimulus-related processes in visual cortex.

The correlation between template level and performance in Experiment 1 raises the question of what an efficient “high-level” attentional template, represented in pTC, might look like. In naturalistic vision, the search target (e.g., a person) is defined by a complex combination of low-level features that vary from one situation to the next because of moment-to-moment differences in viewpoint, lighting, viewing distance, and occlusion, among other factors. Furthermore, search targets are typically

embedded in cluttered visual scenes with a large number of competing objects that heavily overlap with the search target in terms of their low-level features. Therefore, an efficient attentional template would need to be composed of relatively complex features that are invariant to viewpoint and size differences and, at the same time, are relatively unique to the target category. Behavioral studies have provided evidence that category level attentional templates are composed of intermediate level category-diagnostic features such as the wheel/rim of a car (Reeder & Peelen, 2013; Delorme, Richard, & Fabre-Thorpe, 2010; Evans & Treisman, 2005). Results of Experiment 2 of the current study further indicate that different mechanisms underlie search at different levels of specificity, consistent with behavioral studies showing that the contents of the search template depend on task demands and target-distractor similarity (Bravo & Farid, 2009, 2012; Schmidt & Zelinsky, 2009; Yang & Zelinsky, 2009; Vickery, King, & Jiang, 2005). Further characterization of effective and ineffective templates for various naturalistic detection tasks will be important as it may be used to improve object detection, for example, by training individuals to adopt more effective templates.

Results of Experiment 2 provide evidence that pTC is causally involved in preparing for category level search to a greater extent than EVC. These results are consistent with two interpretations regarding the contribution of these regions to category level search. Perhaps, the most straightforward interpretation, also considering the results of Experiment 1, is that pTC stimulation disrupted category level search task performance. Alternatively, or in addition, it is possible that EVC stimulation facilitated category level search task performance, possibly by disrupting inhibitory contributions of EVC (e.g., Mullin & Steeves, 2011). For example, the fMRI study on which the current study was based (Peelen & Kastner, 2011) found both a positive correlation between behavioral performance and categorical information contained in pTC activity patterns and a negative correlation between behavioral performance and categorical information contained in EVC activity patterns.

Experiment 2 did not reveal a significant effect of EVC stimulation (relative to pTC stimulation) on individual level task performance. Such an effect would be predicted if participants used low-level features (e.g., line orientation) to perform the individual level task. It should be noted, however, that even the targets in the individual level task varied considerably in size and location and were placed in complex natural scenes with distractors varying in low-level features. We therefore expect that participants used midlevel features (e.g., line conjunctions, simple shapes) rather than low-level features to solve this task. These features are likely represented in regions not targeted by TMS in our study, perhaps in extrastriate areas located between EVC and pTC.

The finding that TMS before scene onset influenced the accuracy of object detection in scenes highlights



the importance of preparatory brain states to visual perception. Our results add to a body of literature showing that prestimulus TMS can affect stimulus detection. For example, TMS over EVC before stimulus onset has been shown to reduce stimulus visibility (Jacobs, Goebel, & Sack, 2012). Other studies have found that motion detection is impaired when TMS is applied over motion-selective V5/MT before stimulus onset (Stevens, McGraw, Ledgeway, & Schluppeck, 2009; Laycock, Crewther, Fitzgerald, & Crewther, 2007). Prestimulus TMS over V5/MT has also been shown to disrupt predictions based on apparent motion (Vetter, Grosbras, & Muckli, 2015). Finally, rhythmic TMS at specific frequencies before stimulus onset affects stimulus visibility in a retinotopically specific manner (Romei, Gross, & Thut, 2010).

Our study differs from these studies in that we disrupted internally maintained representations of top-down attentional sets rather than spontaneous ongoing activity. The maintenance of a search template is closely related to the maintenance of items in STM, with both processes biasing attention to stimuli that match currently active representations (Gazzaley & Nobre, 2012). Previous studies have shown that TMS over visual cortex during the retention interval (i.e., in the absence of visual stimulation) can interfere with active memory representations (Zokaei, Manohar, Husain, & Feredoes, 2014; van de Ven & Sack, 2013) and modulate subsequent attentional selection (Soto, Llewelyn, & Silvanto, 2012). Our results extend these previous findings by providing evidence that TMS can disrupt the top-down attentional set for categorical targets in real-world scenes.

Attention mechanisms have developed and evolved to optimally select task-relevant objects in real-world scenes, yet these mechanisms have only recently become the topic of neuroscientific study. The current results point to an important role for internally generated preparatory activity in visual cortex in mediating the detection of object categories in cluttered daily-life environments.

## Acknowledgments

The research was funded by the Autonomous Province of Trento, Call “Grandi Progetti 2012,” Project “Characterizing and improving brain mechanisms of attention—ATTEND.”

Reprint requests should be sent to Marius V. Peelen, Center for Mind/Brain Sciences, University of Trento, Corso Bettini 31, 38068 Rovereto (TN), Italy, or via e-mail: marius.peelen@unitn.it.

## REFERENCES

- Bansal, A. K., Madhavan, R., Agam, Y., Golby, A., Madsen, J. R., & Kreiman, G. (2014). Neural dynamics underlying target detection in the human brain. *Journal of Neuroscience*, *34*, 3042–3055.
- Bravo, M. J., & Farid, H. (2009). The specificity of the search template. *Journal of Vision*, *9*, 34.1–34.9.
- Bravo, M. J., & Farid, H. (2012). Task demands determine the specificity of the search template. *Attention, Perception, & Psychophysics*, *74*, 124–131.
- Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, *363*, 345–347.
- Delorme, A., Richard, G., & Fabre-Thorpe, M. (2010). Key visual features for rapid categorization of animals in natural scenes. *Frontiers in Psychology*, *1*, 21.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- Downing, P. E., Wiggett, A. J., & Peelen, M. V. (2007). Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *The Journal of Neuroscience*, *27*, 226–233.
- Duecker, F., & Sack, A. T. (2013). Pre-stimulus sham TMS facilitates target detection. *PLoS One*, *8*, e57765.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.
- Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1476–1492.
- Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*, 129–135.
- Jacobs, C., Goebel, R., & Sack, A. T. (2012). Visual awareness suppression by pre-stimulus brain stimulation: A neural effect. *Neuroimage*, *59*, 616–624.
- Koivisto, M., Railo, H., Revonsuo, A., Vanni, S., & Salminen-Vaparanta, N. (2011). Recurrent processing in V1/V2 contributes to categorization of natural scenes. *Journal of Neuroscience*, *31*, 2488–2492.
- Larsson, J., & Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *The Journal of Neuroscience*, *26*, 13128–13142.
- Laycock, R., Crewther, D. P., Fitzgerald, P. B., & Crewther, S. G. (2007). Evidence for fast signals and later processing in human V1/V2 and V5/MT+: A TMS study of motion perception. *Journal of Neurophysiology*, *98*, 1253–1262.
- Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, *16*, 5154–5167.
- Mullin, C. R., & Steeves, J. K. (2011). TMS to the lateral occipital cortex disrupts object processing but facilitates scene processing. *Journal of Cognitive Neuroscience*, *23*, 4174–4184.
- Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, *460*, 94–97.
- Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *108*, 12125–12130.
- Peelen, M. V., & Kastner, S. (2014). Attention in the real world: Toward understanding its neural basis. *Trends in Cognitive Sciences*, *18*, 242–250.
- Reeder, R. R., & Peelen, M. V. (2013). The contents of the search template for category-level search in natural scenes. *Journal of Vision*, *13*, 13.
- Romei, V., Gross, J., & Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: Correlation or causation? *Journal of Neuroscience*, *30*, 8692–8697.
- Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (2008). LabelMe: A database and web-based tool for image annotation. *International Journal of Computer Vision*, *77*, 157–173.
- Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue.



- Quarterly Journal of Experimental Psychology*, 62, 1904–1914.
- Schwarzbach, J. (2011). A simple framework (ASF) for behavioral and neuroimaging experiments based on the psychophysics toolbox for MATLAB. *Behavior Research Methods*, 43, 1194–1201.
- Soto, D., Llewelyn, D., & Silvanto, J. (2012). Distinct causal mechanisms of attentional guidance by working memory and repetition priming in early visual cortex. *Journal of Neuroscience*, 32, 3447–3452.
- Stevens, L. K., McGraw, P. V., Ledgeway, T., & Schluppeck, D. (2009). Temporal characteristics of global motion processing revealed by transcranial magnetic stimulation. *European Journal of Neuroscience*, 30, 2415–2426.
- van de Ven, V., & Sack, A. T. (2013). Transcranial magnetic stimulation of visual cortex in memory: Cortical state, interference and reactivation of visual content in memory. *Behavioural Brain Research*, 236, 67–77.
- Vetter, P., Grosbras, M. H., & Muckli, L. (2015). TMS over V5 disrupts motion prediction. *Cerebral Cortex*, 25, 1052–1059.
- Vickery, T. J., King, L. W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, 5, 81–92.
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 419–433.
- Wolfe, J. M., Vo, M. L., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, 15, 77–84.
- Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically-defined targets. *Vision Research*, 49, 2095–2103.
- Zokaei, N., Manohar, S., Husain, M., & Feredoes, E. (2014). Causal evidence for a privileged working memory state in early visual cortex. *Journal of Neuroscience*, 34, 158–162.

Uncorrected Proof

## AUTHOR QUERY

### **AUTHOR PLEASE ANSWER QUERY**

During the preparation of your manuscript, the question listed below arose. Kindly supply the necessary information.

1. Please spell out EBA.

**END OF QUERY**

Uncorrected Proof