**Preparatory activity in posterior temporal cortex causally
contributes to object detection in scenes**

Journal:	<i>Journal of Cognitive Neuroscience</i>
Manuscript ID:	JOCN-2015-0074.R1
Manuscript Type:	Original
Date Submitted by the Author:	n/a
Complete List of Authors:	Reeder, Reshanne; University of Trento, CIMEC Perini, Francesca; University of Trento, CIMEC Peelen, Marius; University of Trento, CIMEC
Keywords:	Attention: Visual, Perception: High-level vision, Transcranial magnetic stimulation

**Preparatory activity in posterior temporal cortex causally contributes to
object detection in scenes**

Reshanne R. Reeder^{a,b}, Francesca Perini^a, Marius V. Peelen^a

^aCenter for Mind/Brain Sciences (CIMEC), University of Trento, 38068 Rovereto
(TN), Italy

^bOtto-von-Guericke University, Institute of Psychology, 39106 Magdeburg, Germany

Corresponding author:

Marius V. Peelen

Center for Mind/Brain Sciences, University of Trento

Corso Bettini 31

38068 Rovereto (TN), Italy

phone +39 0464 808718

e-mail: marius.peelen@unitn.it

Acknowledgements

The research was funded by the Autonomous Province of Trento, Call "Grandi
Progetti 2012", project "Characterizing and improving brain mechanisms of
attention - ATTEND".

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 the present 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 temporal cortex (pTC) identified by fMRI as carrying category-specific preparatory activity patterns. Results showed that TMS applied over pTC prior to scene onset (-200 ms, -100 ms) 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 prior to 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 and Humphreys, 1989, Wolfe et al., 1989), with the degree to which an object matches this internal description influencing the likelihood that the object is attended.

While the concept of an attentional template is intuitively appealing, its specific neural implementation and functional significance remain unclear, particularly in naturalistic vision (Peelen and Kastner, 2014). One hypothesis is that preparing to detect a target stimulus results in the pre-activation or priming of neurons that represent that stimulus in visual cortex (Desimone and Duncan, 1995). This pre-activation 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 pre-stimulus increases in target-selective populations of neurons and voxels (Chelazzi et al., 1993, Peelen & Kastner, 2011). 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 of visual cortex, for example in regions of prefrontal cortex implicated in working memory (Miller et al.,

1996, Bansal et al., 2014), with attentional selection mediated by feedback from these regions only after the onset of a search display.

In the present 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 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

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

onset (“pre-scene TMS”) negatively affects detection performance relative to pre-scene 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 pre-scene TMS, testing whether pre-scene TMS over pTC, relative to pre-scene TMS over EVC, selectively affects detection at the category level relative to detection based on lower-level features.

For Review Only

Materials and Methods

Participants. A total of 52 healthy right-handed undergraduate and graduate students from the University of Trento participated in the study. Sixteen of these participated in Experiment 1 (8 female, aged 21-31 years, mean=25.5 years) and 40 (4 who also participated in Experiment 1) participated in Experiment 2 (25 female, 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 standard deviations 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 prior to 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 prior to 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, subjects sat in a straight-backed chair and rested their heads on a chinrest 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 arm. TMS pulses were delivered with a

Magstim Rapid stimulator with a 75mm MCF-B65 Butterfly coil (Experiment 1) and a Magstim Rapid stimulator with a 50mm 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 subject based on the minimum stimulation intensity needed to produce a visible right hand twitch in at least 5/10 pulses. Stimulation intensity for experimentation was set at 120% of each subject's MT. Motor threshold, 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: xyz = 46, -58, 8); see Figure 1 for an illustration of the pTC coordinates mapped on a sample brain. This region is located in posterior temporal cortex, about 2-3 cm anterior to retinotopically-defined regions L01 ($y = -89$) and L02 ($y = -82$; Larsson & Heeger, 2006), and about 1 cm anterior to the peaks of object-selective L0 ($y = -71$), body-selective EBA ($y = -67$), and motion-selective MT+ ($y = -65$; Downing et al., 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

1
2
3 produce inferior-to-superior current flow to find the best stimulation angle that
4
5 minimized muscle twitching and eye blinking.
6
7

8 In Experiment 1, effects of pTC stimulation were compared with effects of
9
10 vertex stimulation. Vertex was located by placing the coil centered between the two
11
12 cerebral hemispheres on the top of the scalp halfway between the inion and the
13
14 nasion. Single-pulse TMS was administered at one of five time points relative to
15
16 scene onset (-200 ms, -100 ms, 0 ms, +100 ms, +200 ms). Stimulation times were
17
18 randomized with an equal number of trials for each of the stimulation times within a
19
20 block. In Experiment 2, effects of pTC stimulation were compared with effects of
21
22 early visual cortex (EVC) stimulation. EVC was located by placing the coil 2 cm
23
24 above the inion pointed inferiorly and adjusting it slightly to the left and right until
25
26 stimulation evoked the perception of phosphenes (flashes of light) in blindfolded
27
28 participants. Stimulation intensity was varied until participants reported seeing
29
30 phosphenes. For participants who did not report seeing phosphenes (15/40
31
32 participants), EVC was located by placing the coil 2cm above the inion pointed
33
34 inferiorly. Double-pulse TMS was administered starting 200 ms prior to scene onset
35
36 (i.e., pulses were delivered at -200 ms and -100 ms relative to scene onset) on half of
37
38 trials (randomly selected). No TMS was delivered on the other half of trials to
39
40 prevent the coil from overheating during an experimental session. No significant
41
42 accuracy differences between conditions were found on trials without TMS ($p >$
43
44 0.18, for all tests). Overall performance was faster for TMS (703 ms) than for no-
45
46 TMS trials (736 ms; $t(39) = 6.79, p < 0.001$, two-tailed), likely reflecting an alerting
47
48 or arousal effect due to the clicking sound that accompanies the pre-scene TMS
49
50
51
52
53
54
55
56
57
58
59
60

pulses. Perhaps due to this alerting effect, accuracy performance was also slightly, though reliably higher for TMS (86.1%) than for no-TMS trials (85.2%; $t(39) = 2.19$, $p = 0.035$, 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-inch gamma-corrected DELL 1908FP-BLK monitor with a screen resolution of 1280x960 pixels and a 60Hz refresh rate (Experiment 1), and a 22-inch Dell E228 WFP monitor with a screen resolution of 1680x1050 pixels and 60Hz 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 et al., 2009, Peelen & Kastner, 2011). 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 (0.81° visual angle) for 1000 ms, followed by a centrally presented single letter cue (1.83° visual angle) for

500 ms, 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 s, followed by a centrally presented photograph of a real-world scene (16.08 x 12.18° visual angle) for 83 ms. All four scene types were presented an equal 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 ms, after which a final fixation cross appeared for 2 s, for a total trial duration of approximately 4.9 seconds. 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 8 experimental blocks of 80 trials each.

TMS was applied over pTC or vertex alternating every two blocks (e.g., 2 blocks pTC stimulation, 2 blocks vertex stimulation, 2 blocks pTC, etc.). The order in which pTC and vertex were stimulated was counterbalanced across participants.

Following 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 et al., 2008), converted to grayscale. 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 a total of 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 (0.81° visual angle) for 1000 ms, followed by a centrally presented single letter cue (1.83° visual angle) for 500 ms, 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 s, followed by a centrally presented photograph of a real-world scene ($16.08 \times 12.18^\circ$ visual angle) for 83 ms in the category-level blocks or 50 ms 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”, “M”) appeared an equal number of times in each of the different scene types. Scenes were backward masked for 350 ms, after which a final fixation cross appeared for 2 s, for a total trial duration of approximately 4.9 seconds. 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, while 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 6 consecutive category-level blocks and 6 consecutive individual-level blocks, with 3 consecutive blocks of pTC stimulation and 3 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 a total of 10 subjects 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

1
2
3 stimulation in the individual-level condition, and 4=EVC stimulation in the
4
5 individual-level condition, then there were 10 participants each in the following task
6
7 orders: 1-2-3-4, 2-1-4-3, 3-4-1-2, and 4-3-2-1). Because the first 20 participants
8
9 were tested with 2 of the 4 possible condition orders, we tested another 20
10
11 participants to complete the full counterbalancing.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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 *Materials and 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 *Materials and Methods*). The critical trials were those in which TMS was applied before scene onset (-200 ms, -100 ms), 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 ms before scene onset ($t(14) = -2.69, p = 0.009$, one-tailed; Figure 4). The corresponding effect at -200 ms approached significance ($t(14) = -1.64, p = 0.06$, one-tailed). There were no significant differences between pTC and vertex when TMS was applied at or after scene onset (0 ms, +100 ms, +200 ms; $p > 0.3$, for all tests). Pre-scene TMS over pTC, averaged across the -200 ms and -100 ms stimulation onsets (shaded area in Figure 4), significantly reduced accuracy relative to pre-scene TMS over vertex ($t(14) = -3.03, p = 0.0045$, one-tailed).

Analysis of RT revealed no significant differences between pTC and vertex stimulation for any of the onset conditions ($p > 0.26$, for all tests).

Next, we turned to the questionnaire data to test our second hypothesis: that the effect of pre-scene TMS (averaged across the -200 ms and -100 ms 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 pre-scene TMS over pTC (relative to pre-scene TMS over vertex) was negatively correlated with template level ($r(13) = -0.576$, $p = 0.012$, 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 computed 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 ($r(13) = -0.632$, $p = 0.006$, one-tailed).

Analysis of RT data showed no significant relationship between the effect of pre-scene TMS over pTC (relative to vertex) and template level ($r(13) = 0.161$, $p = 0.58$, two-tailed).

Experiment 2

Experiment 2 was conducted to test the hypothesis that pre-scene TMS over pTC, relative to pre-scene 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

1
2
3 templates. By contrast, in the individual-level condition the same person or car was
4
5 presented throughout a block, allowing for detection based on lower-level features.
6
7 In behavioral pilots, we found that the individual-level condition was indeed easier
8
9 than the category-level condition, indicating that participants made use of the
10
11 predictable features to optimize their performance. Based on these pilot data we
12
13 adjusted the presentation times to equate the difficulty of the two tasks. The two
14
15 tasks were matched by presenting scenes for 83 ms in the category-level task and
16
17 for 50 ms in the individual-level task (category-level task: 86.4%, individual level
18
19 task: 85.9%; $F(1,39) = 0.34, p = 0.56$, two-tailed). Double-pulse TMS was delivered
20
21 before scene onset (-200 ms and -100 ms) to disrupt preparatory processing in pTC
22
23 or EVC, in different blocks.
24
25
26
27
28
29

30 Confirming our hypothesis and replicating Experiment 1, pre-scene TMS over
31
32 pTC reduced accuracy on the category-level detection task, relative to pre-scene
33
34 TMS over EVC ($t(39) = -2.34, p = 0.012$, one-tailed; Figure 5). By contrast, for the
35
36 individual-level task, pre-scene TMS over pTC had no effect ($t(39) = -0.05, p = 0.96$),
37
38 resulting in a significant Task (category-level, individual-level) x Region (pTC, EVC)
39
40 interaction ($F(1,39) = 3.2, p = 0.041$, one-tailed).
41
42
43

44 Analysis of RT data confirmed these results, showing slower responses in the
45
46 category-level task when TMS was applied over pTC relative to EVC ($t(39) = 2.93, p$
47
48 $= 0.003$, one-tailed). TMS had the opposite effect in the individual-level task, with
49
50 TMS over pTC speeding up responses relative to TMS over EVC ($t(39) = -2.08, p =$
51
52 0.022 , one-tailed); the Task (category, individual) x Region (pTC, EVC) interaction
53
54 was significant ($F(1,39) = 14.6, p < 0.001$, one-tailed).
55
56
57
58
59
60

Discussion

Theoretical and empirical studies have suggested an important role for top-down attentional templates in guiding real-world attentional selection (for reviews, see Wolfe et al., 2011, Peelen & Kastner, 2014), but the neural implementation and functional significance of such templates have remained elusive. In the present 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 just 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 pre-scene 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

1
2
3 in category-level search. The finding that pTC represents relatively high-level object
4 representations is in accordance with its position in the visual system hierarchy.
5
6 The findings of a correlation between template level and the TMS effect (Experiment
7 1) and the specificity of TMS to category-level detection tasks (Experiment 2) are
8 important in that they rule out the possibility that the performance impairments
9 following TMS over pTC were due to TMS-related artifacts (e.g., eye blinks or muscle
10 twitches).
11
12
13
14
15
16
17
18
19

20 In Experiment 1, TMS over pTC *after* scene onset did not significantly
21 interfere with performance. This could indicate a dissociation between the
22 maintenance of a top-down attentional set (critically involving pTC) and the visual
23 analysis of the scene itself (not critically involving pTC). However, there may be
24 several other reasons for the absence of a post-scene TMS effect. Perhaps most
25 importantly, the specific times at which TMS was delivered relative to scene onset in
26 the current study (0 ms, +100 ms, and +200 ms) may have missed a critical time
27 window at which this region is involved in visual processing. For example, a recent
28 study investigating animal detection in scenes found that TMS over lateral occipital
29 cortex affected subjective perception when applied 150 ms after scene onset, but
30 not when applied 90 ms or 210 ms after scene onset (Koivisto et al., 2011). In
31 contrast, the critical time window of pre-scene TMS for disrupting preparatory
32 templates would be expected to be broader than this. Another possibility for the
33 lack of a significant post-scene TMS effect is that the effect of stimulation to pTC was
34 simply too weak to disrupt visually driven processing, with internally generated
35 preparatory activity being more susceptible to interference. More work is needed to
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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 due to 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 comprised 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 (Evans and Treisman, 2005, Delorme et al., 2010, Reeder & Peelen, 2013). 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-distracter similarity (Bravo & Farid, 2009; 2012; Schmidt & Zelinsky, 2009, Vickery et al., 2005, Yang & Zelinsky, 2009). 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 additionally, 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 mid-level 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 early visual cortex and posterior temporal cortex.

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 pre-stimulus TMS can affect stimulus detection. For example, TMS over early visual cortex (EVC) before stimulus onset has been shown to reduce stimulus visibility (Jacobs et al., 2012). Other studies have found that motion detection is impaired when TMS is applied over motion-selective V5/MT before stimulus onset (Laycock et al., 2007, Stevens et al., 2009). Pre-stimulus TMS over V5/MT has also been shown to disrupt predictions based on apparent motion (Vetter et al., 2013). Finally, rhythmic TMS at specific frequencies before stimulus onset affects stimulus visibility in a retinotopically specific manner (Romei et al., 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 short-term memory, 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 et al., 2014; van de Ven & Sack, 2013), and modulate subsequent attentional selection (Soto et al., 2012). Our results extend these

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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.

For Review Only

References

- Bansal AK, Madhavan R, Agam Y, Golby A, Madsen JR, Kreiman G (2014) Neural dynamics underlying target detection in the human brain. *J Neurosci* 34:3042-3055.
- Bravo MJ, Farid H (2009) The specificity of the search template. *J Vis* 9:34 31-39.
- Bravo MJ, Farid H (2012) Task demands determine the specificity of the search template. *Atten Percept Psychophys* 74:124-131.
- Chelazzi L, Miller EK, 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. *Front Psychol* 1:21.
- Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193-222.
- Downing, PE, Wiggett, AJ, & Peelen, MV (2007). Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *The Journal of neuroscience*, 27(1), 226-233.
- Duecker, F, & Sack, AT (2013). Pre-stimulus sham TMS facilitates target detection. *PloS one*, 8(3), e57765.
- Duncan J, Humphreys GW (1989) Visual search and stimulus similarity. *Psychol Rev* 96:433-458.
- Evans KK, Treisman A (2005) Perception of objects in natural scenes: is it really attention free? *J Exp Psychol Hum Percept Perform* 31:1476-1492.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Gazzaley A, Nobre AC (2012) Top-down modulation: bridging selective attention and working memory. *Trends Cogn Sci* 16:129-135.

Jacobs C, Goebel R, Sack AT (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. *J Neurosci* 31:2488-2492.

Larsson, J, & Heeger, DJ (2006). Two retinotopic visual areas in human lateral occipital cortex. *The Journal of Neuroscience*, 26(51), 13128-13142.

Laycock R, Crewther DP, Fitzgerald PB, Crewther SG (2007) Evidence for fast signals and later processing in human V1/V2 and V5/MT+: A TMS study of motion perception. *J Neurophysiol* 98:1253-1262.

Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16:5154-5167.

Mullin CR, Steeves JK (2011) TMS to the lateral occipital cortex disrupts object processing but facilitates scene processing. *J Cogn Neurosci* 23:4174-4184.

Peelen MV, Fei-Fei L, Kastner S (2009) Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460:94-97.

Peelen MV, Kastner S (2011) A neural basis for real-world visual search in human occipitotemporal cortex. *Proc Natl Acad Sci U S A* 108:12125-12130.

Peelen MV, Kastner S (2014) Attention in the real world: toward understanding its neural basis. *Trends Cogn Sci* 18:242-250.

- 1
2
3 Reeder RR, Peelen MV (2013) The contents of the search template for category-level
4
5 search in natural scenes. *J Vis* 13:13.
6
7
8 Romei V, Gross J, Thut G (2010) On the role of prestimulus alpha rhythms over
9
10 occipito-parietal areas in visual input regulation: correlation or causation? *J*
11
12 *Neurosci* 30:8692-8697.
13
14
15 Russell BC, Torralba A, Murphy KP, Freeman WT (2008) LabelMe: A database and
16
17 web-based tool for image annotation. *International Journal of Computer*
18
19 *Vision* 77:157-173.
20
21
22 Schmidt J, Zelinsky GJ (2009) Search guidance is proportional to the categorical
23
24 specificity of a target cue. *Q J Exp Psychol (Hove)* 62:1904-1914.
25
26
27 Schwarzbach J (2011) A simple framework (ASF) for behavioral and neuroimaging
28
29 experiments based on the psychophysics toolbox for MATLAB. *Behav Res*
30
31 *Methods* 43:1194-1201.
32
33
34 Soto D, Llewelyn D, Silvanto J (2012) Distinct causal mechanisms of attentional
35
36 guidance by working memory and repetition priming in early visual cortex. *J*
37
38 *Neurosci* 32:3447-3452.
39
40
41 Stevens LK, McGraw PV, Ledgeway T, Schluppeck D (2009) Temporal characteristics
42
43 of global motion processing revealed by transcranial magnetic stimulation.
44
45 *Eur J Neurosci* 30:2415-2426.
46
47
48 van de Ven V, Sack AT (2013) Transcranial magnetic stimulation of visual cortex in
49
50 memory: cortical state, interference and reactivation of visual content in
51
52 memory. *Behav Brain Res* 236:67-77.
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Vetter P, Grosbras MH, Muckli L (2013) TMS Over V5 Disrupts Motion Prediction.
Cereb Cortex.

Vickery TJ, King LW, Jiang Y (2005) Setting up the target template in visual search. J
Vis 5:81-92.

Wolfe JM, Cave KR, Franzel SL (1989) Guided search: an alternative to the feature
integration model for visual search. J Exp Psychol Hum Percept Perform
15:419-433.

Wolfe JM, Vo ML, Evans KK, Greene MR (2011) Visual search in scenes involves
selective and nonselective pathways. Trends Cogn Sci 15:77-84.

Yang H, Zelinsky GJ (2009) Visual search is guided to categorically-defined targets.
Vision Res 49:2095-2103.

Zokaei N, Manohar S, Husain M, Feredoes E (2014) Causal evidence for a privileged
working memory state in early visual cortex. J Neurosci 34:158-162.

Figure Legends

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

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 5 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, while 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 ms for the individual-level task and 83 ms for the category-level task. Double-pulse TMS was applied over pTC or EVC before scene onset (-200 ms, -100 ms).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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. Pre-scene TMS over pTC, relative to vertex, significantly reduced detection accuracy. Error bars indicate SEM.

Figure 5. Results of Experiment 2, showing the effect of pre-scene TMS (double pulse at- 200 ms and -100 ms) 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.

Figure 1. Region in posterior temporal cortex (pTC) stimulated with TMS in Experiments 1 and 2. pTC location was based on fMRI results of Peelen & Kastner (2011).
116x72mm (300 x 300 DPI)

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 5 time points relative to scene onset.

239x178mm (300 x 300 DPI)

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, while 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 ms for the individual-level task and 83 ms for the category-level task. Double-pulse TMS was applied over pTC or EVC before scene onset (-200 ms, -100 ms).
238x171mm (300 x 300 DPI)

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. Pre-scene TMS over pTC, relative to vertex, significantly reduced detection accuracy. Error bars indicate SEM.
209x122mm (300 x 300 DPI)

Figure 5. Results of Experiment 2, showing the effect of pre-scene TMS (double pulse at- 200 ms and -100 ms) 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.
116x93mm (300 x 300 DPI)