Journal of Experimental Psychology:
Human Perception and Performance
2014, Vol. 40, No. 1, 50–60

0096-1523/14/$12.00 DOI: 10.1037/a0033228

Rapid Guidance of Visual Search by Object Categories

Rebecca Nako
University of London

Rachel Wu
University of Rochester

Martin Eimer
University of London

Visual search is often controlled by attentional templates that represent specific target items or target features, but can also be directed toward object categories. We studied the relationship between item-based and category-guided attentional control during visual search for one specific item (e.g., the letter *C*), two or three items (e.g., the letters *C, F,* and *X*), or categorically defined targets (e.g., any letter). To assess the efficiency of visual search for single, multiple, or category-defined targets, we measured the N2pc component as an electrophysiological marker of attentional target selection. In Experiment 1, where targets were presented among distractors from a different category (e.g., letters among digits), a category-based selection strategy was available. Category-based attentional control triggered spatially selective modulations of visual–perceptual processing that emerged within less than 200 ms after stimulus onset and preceded the effects of item-specific attentional templates. In Experiment 2, where letter targets appeared among letter distractors, target detection could no longer be guided by categorical top-down task sets. Search efficiency decreased as the target set size increased, in line with capacity limitations for item-specific attentional templates. Results demonstrate that category-based attentional guidance can be used rapidly and efficiently during visual search for alphanumeric targets.

Keywords: attentional selection, N2pc, visual search, category search

Before going to work in the morning, you have to look for a number of related items—your keys, phone, wallet, and so forth. How does searching for one item compare with search for multiple items and to search for a category of items (e.g., anything edible for lunch)? Previous work suggests that goal-directed visual search is controlled in a top-down fashion by attentional templates. These templates are active visual working memory representations that specify physical properties of relevant target objects (e.g., Desimone & Duncan, 1995; Olivers, Peters, Houtkamp, & Roelfsema, 2011). Attentional templates can represent visual features (e.g., color, orientation, or shape) or whole objects (e.g., apples, cats, or teddy bears). Activated templates bias cognitive and neural processing in favor of objects that match the current target-defining attributes (Desimone & Duncan, 1995).

This article was published Online First June 24, 2013.

Rebecca Nako, Department of Psychological Sciences, Birkbeck, University of London, London, England; Rachel Wu, Department of Brain and Cognitive Sciences, University of Rochester; Martin Eimer, Department of Psychological Sciences, Birkbeck, University of London.

Rebecca Nako and Rachel Wu made equal contributions to this article.

This research was funded by a British Academy grant (SG- 110780) awarded to Rachel Wu, Martin Eimer, Gaia Scerif, and Richard Aslin. We thank Susan Nicholas for technical support and Gaia Scerif, Richard Aslin, and Jenny Roche for helpful discussions and comments on earlier versions of this article. We also thank Susanne Ferber, Clayton Hickey, and one anonymous reviewer for insightful comments and feedback on this work.

Correspondence concerning this article should be addressed to Rachel Wu, Brain and Cognitive Sciences, Meliora Hall Box 270268, University of Rochester, Rochester, NY 14627. E-mail: rwu@bcs.rochester.edu

When searching for objects, attentional templates can be defined by a specific feature (e.g., color) or by a distinct combination of features (conjunctive search; e.g., color and shape). Search for unrelated physically dissimilar complex objects (e.g., your keys and your wallet) requires guidance by simultaneously active templates that match each objects' distinctive features. It has been argued that only a single attentional template can be active at any given time (Olivers et al., 2011). In line with this hypothesis, search for multiple dissimilar target features or objects is considerably less efficient than search for one specific feature or object (e.g., Houtkamp & Roelfsema, 2009; Wolfe, 2012; Grubert & Eimer, 2012).

However, target detection cannot always be based on a straightforward physical match between visual features and a specific attentional template. This is the case when search is directed toward an object category, in particular when different members of the same category are perceptually dissimilar (e.g., when you search in the kitchen for something to eat). Because template-based guidance is not available, visual search for category-defined objects should be less efficient than search for one specific target feature or object. Yang and Zelinsky (2009) found that categorical search is slower than search for a specific exemplar of the category, but still quicker than random search. Search efficiency improves when feature information about the target is available (Malcolm & Henderson, 2009) and when there is increased dissimilarity between category exemplars and distractors (Alexander & Zelinsky, 2011) or increased similarity between distractors (Alexander & Zelinsky, 2012). This suggests that visual feature information can play an important role during category-based search, in particular when categories include perceptually similar items.

Even when the attentional selection of category-defined objects cannot be guided by physically matching templates, it still produces spatially selective modulations of visual processing at relatively early perceptual stages in extrastriate visual cortex. This was demonstrated in a recent set of experiments from our lab (Wu et al., 2013) that combined behavioral and electrophysiological measures of attentional object selection during feature-based and category-based visual search. On each trial, a cue specified the target in an upcoming search array either at the item level (e.g., find the letter *C*) or categorically (e.g., find any letter). To assess the time course and efficiency of item-based versus category-based visual search, we measured the N2pc component as an event-related brain potential (ERP) marker of attentional target selection (Luck & Hillyard, 1994; Eimer, 1996; Luck, Girelli, McDermott, & Ford, 1997). The N2pc is an enhanced negativity at occipitotemporal electrodes contralateral to the hemifield of a visual candidate target object that is generated in retinotopic extrastriate occipitotemporal cortex (Hopf et al., 2000) and reflects the spatially selective attentional selection of target objects among distractors in visual search arrays. Previous N2pc studies have demonstrated that this component can be reliably present from as early as 180 ms after stimulus onset (e.g., Eimer & Kiss, 2008; Grubert & Eimer, 2012). The critical finding of our study (Wu et al., 2013) was that an N2pc was triggered by target objects not only during item-based search that could be guided by a physical match with a stored attentional template, but also during category-based search. This was the case for familiar visual categories (letters vs. digits), even when category set size (12 different possible target letters) exceeded working memory capacity for individually represented items and also for novel categories (Chinese characters) that had to be learned during the experimental sessions. The observation that N2pc components were elicited during the selection of category-defined targets demonstrates that category-based attentional guidance can affect relatively early stages of visual–perceptual processing. However, target reaction times (RTs) were faster and the target-elicited N2pc components were larger and emerged earlier when search was item-based rather than when it was purely category-based, indicating that item-based attentional target selection is more efficient than category-guided selection.

Our previous study (Wu et al., 2013) compared search for category-defined targets and search for one specific target item that could be based on a physical match with an object template. However, there are many situations where we search for several different objects simultaneously (e.g., keys and wallet). How does search for multiple objects compare with purely category-based search? Is searching for category-defined targets less efficient than search for two (or three) specific targets? And do categories play a role in the guidance of attention during multiple-object search? Is search for two or more objects from the same category exclusively controlled by independent representations of each target object, or can it also be guided by category-defined search strategies?

In order to answer these questions, Experiment 1 used behavioral and ERP measures to study the efficiency of attentional target selection in different task conditions where participants searched for one specific item (e.g., the letter *A*), two targets (e.g., the letters *G* or *P*), three targets (e.g., the letters *Q, R,* or *U*), or purely category-defined targets (e.g., any letter). One group of participants searched for letter targets among digit distractors, and the other group searched for digits among letters. On target trials, this item was one of the current specified target items among distractors from the opposite category (e.g., the letter *A* among three digits; see Figure 1). Foil trials included one target category-matching item that was not part of the current target set (e.g., the letter *A* during search for *G* or *P*). On nontarget trials, the search

Examples of Search Arrays in Experiments 1 and 2

	Condition	1T	2T	3T	CT	Foil	Distractor only
Experiment 1	Example targets	A	G, P	Q, R, U	All letters	A	A
	Example search array	4 7 2 A	4 7 2 G	4 7 2 R	4 7 2 R	4 7 2 R	4 7 2 5
Experiment 2	Example targets	A	G, P	Q, R, U	n/a	n/a	A
	Example search array	R G U A	R G U A	R G P A	n/a	n/a	R G P U

Figure 1. (Top) Search stimuli used in Experiments 1 and 2. (Bottom) Examples of targets and search arrays used in the different task conditions in Experiments 1 and 2. Participants had to search for one specific target (1T), two targets (2T), three targets (3T), or category-defined targets (CT).

array contained four irrelevant-category items (four digits during letter search, or vice versa). At the beginning of each experimental block, participants were instructed to search for one target (1T), two targets (2T), three targets (3T), or category-defined targets (CT), and these target specifications remained constant for five successive blocks. Their task was to make a target present/absent decision on each trial.

The efficiency of attentional target selection in these four task conditions was assessed by measuring behavioral performance and the N2pc component on target and foil trials. Search should be most efficient when it can be guided by an item-specific attentional template. Thus, targets should be detected faster, and the target N2pc should be largest and/or emerge earliest during search for one specific target item (Condition 1T). However, and in line with our previous results (Wu et al., 2013), an N2pc should also be triggered by purely category-defined targets (Condition CT), thereby confirming that category-based search can modulate relatively early perceptual stages of visual processing. The critical question concerned the guidance of attentional object selection during search for two or three possible target items (Conditions 2T and 3T) and its differences relative to purely category-based search (Condition CT). If multiple-object search is controlled by independent representations of each target object, target selection efficiency should be impaired in the 2T condition relative to single-item search (where a unique attentional template can be used; Olivers et al., 2011) and even more so in the 3T condition. These differences should be reflected by corresponding behavioral and N2pc differences. More importantly, if search for two or three targets is more efficient than purely category-based search, target RTs should be faster, and target N2pc components earlier and larger in the 2T and 3T conditions relative to the CT condition.

An alternative possibility is that multiple-object search with related objects is not guided by independent representations of each target object, but instead by category-defined search strategies. In this case, there should be no systematic performance and target N2pc differences between the 2T, 3T, and CT task conditions, because target selection is controlled by a category-defined task set in all of these conditions. To assess this possibility, the analysis of N2pc components on foil trials is particularly relevant: If participants used a category-based search strategy (i.e., find any letter) even when they were explicitly asked to search for two or three letters, category-matching items that are not part of the currently active task set should attract attention and thus trigger a reliable N2pc. To investigate the time course of item-based versus object-based attentional guidance, we specifically focused on the initial phase of the N2pc between 180 and 200 ms after search array onset: If item-based control precedes category-based control, a reliable N2pc should be triggered exclusively by targets, but not by foils during this early stage of attentional object selection.

Experiment 1

Methods

Participants. Twenty-one paid volunteers participated in this experiment. Five participants were excluded due to excessive eye movements (more than 47% of all trials removed due to eye movement artifacts). All remaining 16 participants ($M = 29.1$ years old, $SD = 4.77$, range: 19–38 years, seven men) had normal or corrected vision.

Stimuli, design, and procedure. Stimuli were presented on a 24-inch LCD monitor with a 100-Hz refresh rate at a viewing distance of 100 cm. On each trial, a search array consisting of four different items was presented. They were drawn from sets of six possible letters and six possible digits. The digits *2, 4, 5, 6, 7,* and *8* and the uppercase letters *A, G, P, Q, R,* and *U* were selected from the available set of letters and digits in order to match the items from both categories in terms of low-level visual features such as curvature, closure, the presence of intersections and diagonal elements, or symmetry (see Figure 1). Each stimulus was individually generated and adjusted in Adobe Photoshop to ensure that these features were matched between the two categories. The four array elements were arranged at equidistant positions around a central fixation dot at a radial distance of 2° visual angle, as measured from the fixation to the center of each stimulus. Each item subtended 1.72° × 1.72°. All stimuli (letters and digits) were black and were presented against a gray background (luminance: 39.4 cd/m^2). Each search array was presented for 200 ms, followed by an empty intertrial interval of 1400 ms. A central fixation point was continuously present, and participants were instructed to maintain central fixation throughout each experimental block.

The participants' task was to detect and report the presence or absence of a target item (a letter target among three digit distractors, or a digit target among three letter distractors). Letters served as targets and digits as distractors for eight participants, and this assignment was reversed for the other eight participants. Participants were instructed to indicate the presence or absence of a target on each trial by pressing one of two response keys with their right index or middle finger. There were four experimental conditions with different target specifications, as illustrated in Figure 1. In one-target (1T) blocks, participants searched for a single prespecified target (e.g., *A*). In two-targets (2T) and three-targets (3T) blocks, they searched for two (e.g., *G,* P) or three (e.g., *Q, R,* and *U*) possible targets. In Category Targets (CT) blocks, they were instructed to search for any letter (or any digit). Each participant searched for different targets in the 1T, 2T, and 3T conditions (i.e., each of the six items in the category served as target in one of these three conditions). All four conditions also included distractor-only trials where four items from the task-irrelevant category were presented. In the 1T, 2T, and 3T conditions, there were also foil trials with search arrays that contained a target category-matching item (e.g., the letter *R* during search for the letter *A*) among three distractors and required a target-absent response. There were no foil trials in CT blocks, because all items in the target category served as targets in these blocks. All trial types were presented in random order, and the locations of targets and foils (when present) were randomly selected on each trial.

The experiment included 20 blocks with 64 trials per block. Five successive blocks were run for each of the four task conditions, and task order was counterbalanced across participants. Each block in the 1T, 2T, and 3T conditions contained 32 target trials, 16 foil trials, and 16 distractor-only trials. Each block in the CT condition included 32 target trials and 32 distractor-only trials. In all four task conditions, target-present or target-absent responses were required on 50% of all trials, respectively.

EEG recording and data analysis. EEG was DC-recorded from 23 scalp electrodes at standard positions of the extended

10/20 system (500-Hz sampling rate; 40-Hz low-pass filter) against a left-earlobe reference and rereferenced offline to averaged earlobes. The continuous EEG was segmented from −100 to 500 ms relative to the search array onset, and trials were epoched relative to a 100-ms prestimulus baseline. Trials with artifacts (horizontal electro-oculogram exceeding ± 25 μV, vertical electro-oculogram exceeding ± 60 μV, all other channels exceeding ± 80 μV) were removed prior to analysis. Only target trials and foil trials were included in the EEG analyses. Averaged waveforms for trials with correct responses (target-present responses on target trials; target-absent responses on foil trials) were computed for target trials in all four task conditions and for foil trials in the 1T, 2T, and 3T conditions. Separate averages were computed for search arrays with targets/foils on the left and right side, respectively. N2pc amplitudes were quantified on the basis of ERP mean amplitudes obtained between 180 and 300 ms after search array onset at lateral posterior electrodes PO7 and PO8. To investigate the effects of item- and category-based attentional guidance during the earliest stage of attentional object selection, additional analyses were also conducted for N2pc mean amplitudes obtained during the initial ascending flank of this component (180–200-ms poststimulus). Target N2pc onset latencies were compared between task conditions, using the jackknife-based analysis method described by Miller, Patterson, and Ulrich (1998). A 50% peak amplitude criterion was used to define N2pc onset. For pairwise comparisons of experimental effects, Bonferroni corrections were applied where appropriate.

Results

Behavioral results. Figure 2 shows the RTs on correct trials (left panels) and mean accuracy (right panels) for the four different task conditions (1T, 2T, 3T, and CT) on target, foil and distractor-only trials. On target trials, there was a main effect of task condition on RTs, $F(3, 45) = 13.7$, $p < .001$, $\eta^2 = .48$. RTs were faster in 1T blocks (490 ms) compared with 2T, 3T, and CT blocks (542, 552, and 532 ms, respectively, all $p < .001$). There were no reliable RT differences between the 2T, 3T, and CT conditions. A similar RT pattern was present on foil trials, where a main effect of task condition, $F(2, 30) = 33.3$, $p < .001$, $\eta^2 = .69$, was due to

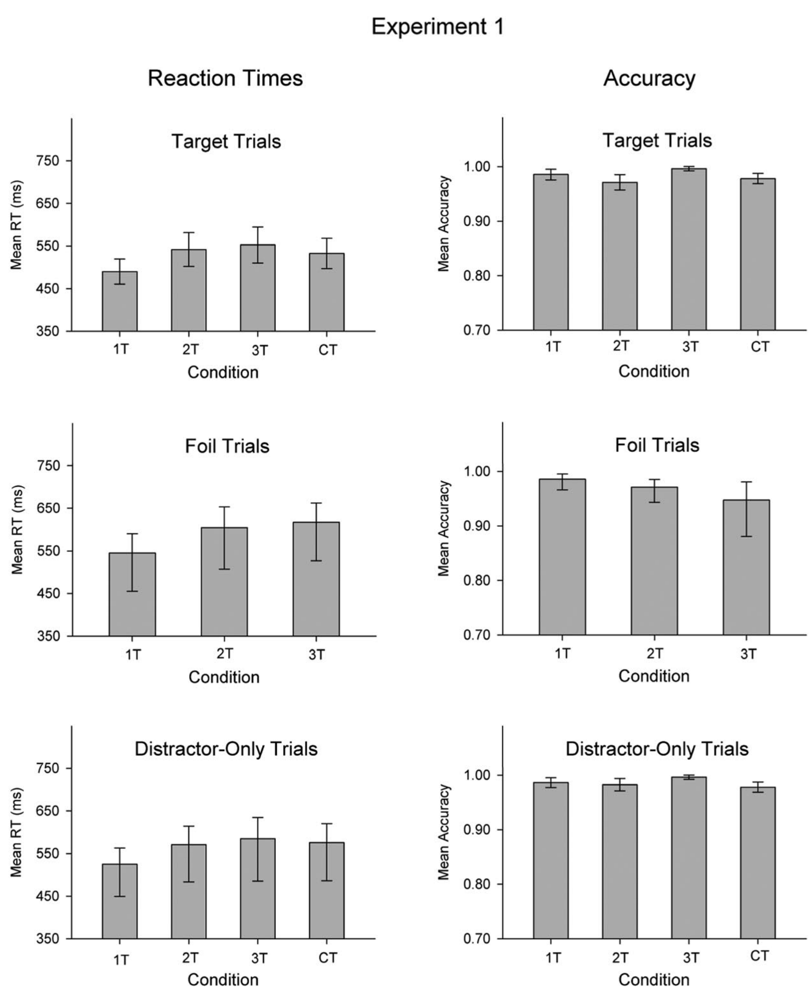

Figure 2. Mean RTs on correct trials (left panels) and mean accuracy (right panels) on target, foil, and distractor-only trials in Experiment 1, shown separately for different task conditions. Error bars indicate ±2 *SE* of mean.

the fact that target-absent responses were faster in 1T blocks (545 ms) compared with 2T and 3T blocks (605 and 617 ms, respectively, both $p < .001$). There was no reliable RT difference between foil trials in the 2T and 3T conditions. Analogous RT differences were also found on distractor-only trials, where a main effect of task condition, $F(3, 45) = 9.0$, $p < .001$, $\eta^2 = .37$, which reflected the fact that target-absent RTs were faster in 1T blocks (525 ms) compared with RTs in 2T, 3T, and CT blocks (570, 585, and 575 ms, all $p \leq .02$, respectively), which did not differ reliably from each other. To test whether target-absent responses were affected by the presence of a foil item, RTs on distractor-only trials and on foil trials were compared with paired *t* tests. Target-absent RTs were reliably slower on nontarget trials that included a foil relative to distractor-only trials, and this was the case not only in Conditions 2T and 3T, but also in 1T blocks; all $t(15) > 3.1$; all $p < .05$.

Accuracy was generally high (see Figure 2, right panels) and did not differ between task conditions for target and for foil trials, $F(3, 45) = 1.9$ and 1.7, $p = .15$ and .20, $\eta^2 = .11$ and .10, respectively. There was a main effect of accuracy on distractor-only trials, $F(3, 45) = 3.98$, $p = .013$, $\eta^2 = .21$. Accuracy was higher in the 3T relative to the CT condition ($p < .05$) and did not differ between any of the other conditions.

ERP results. Figure 3 (left four graphs) shows ERPs triggered on target trials in the 500 ms after search array onset at electrodes PO7/8, separately for all four task conditions (1T, 2T, 3T, CT). Solid and dashed lines show ERPs contralateral and ipsilateral waveforms to the target item. The right graph displays the difference waveforms for each task, obtained by subtracting ipsilateral from contralateral ERPs. Figure 4 (left three graphs) shows ERPs for foil trials in the 1T, 2T, and 3T conditions and the corresponding difference waveforms. Clear N2pc components were triggered on target trials in all four conditions (see Figure 3). Target N2pc amplitudes were larger in 1T blocks relative to the other three types of blocks. N2pc components of similar size were triggered in 2T, 3T, and CT blocks. N2pc components also appeared to be present in response to category-matching foils, although these were reduced in amplitude relative to the target N2pcs (Figure 4, bottom right panel).

Target trials. A repeated-measures ANOVA on ERP mean amplitudes in the N2pc time window (180–300-ms poststimulus) for the factors task condition and laterality (electrode contralateral vs. ipsilateral to the target) revealed a main effect of laterality, $F(1, 15) = 61.1$, $p < .001$, $\eta^2 = .80$, that was accompanied by an interaction between task condition and laterality, $F(3, 45) = 3.2$, $p = .033$, $\eta^2 = .17$. Paired comparisons of contralateral and ipsilateral ERPs confirmed that targets triggered reliable N2pc components in all four task conditions: 1T: $t(15) = 6.86$, $p < .001$; 2T: $t(15) = 6.89$, $p < .001$; 3T: $t(15) = 6.72$, $p < .001$; CT: $t(15) = 8.35$, $p < .001$. A planned Helmert contrast showed that the target N2pc was significantly larger in the 1T condition relative to the other three task conditions, $F(1, 15) = 7.2$, $p = .017$, $\eta^2 = .33$. There were no reliable N2pc amplitude differences between the 2T, 3T, and CT conditions. N2pc onset latencies did not differ reliably between any task condition, F_c (3, 45) = .29, $p = .884$.

Foil trials. N2pc mean amplitudes triggered by foils in the 180–300 ms poststimulus interval were analyzed with a repeated-measures ANOVA for the factors task condition (1T, 2T, 3T) and laterality. There was a main effect of laterality, $F(1, 15) = 93.5$, $p < .001$, $\eta^2 = .86$, reflecting the presence of N2pc components on foil trials. The interaction between task condition and laterality approached significance, $F(2, 30) = 3.1$, $p = .062$, $\eta^2 = .17$, as the foil N2pc tended to be smaller in the 1T condition relative to the 2T and 3T conditions (see Figure 4). Pairwise comparisons of

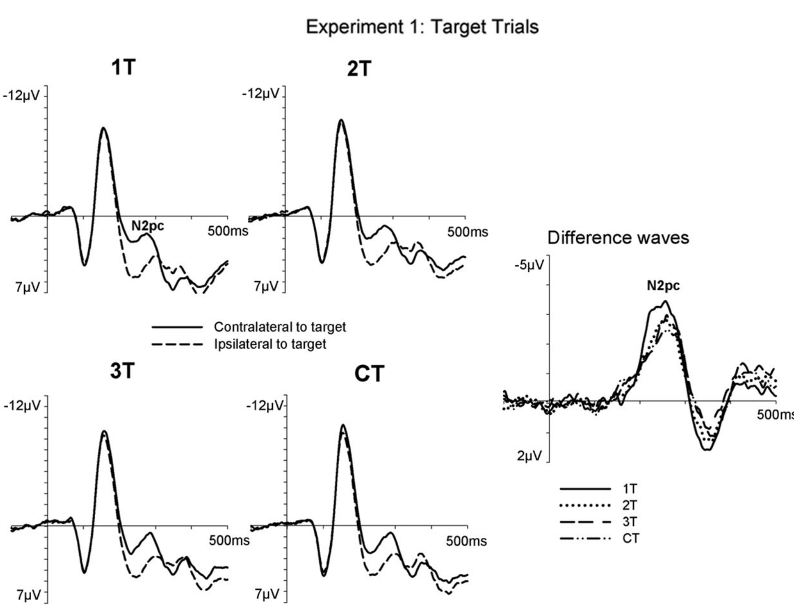

Figure 3. (Left and middle) Grand-average ERPs elicited in Experiment 1 in response to search arrays on target trials at posterior electrodes PO7/8 contralateral and ipsilateral to a target item, shown separately for the four task conditions. (Right) N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERP waveforms at PO7/8 for each task condition.

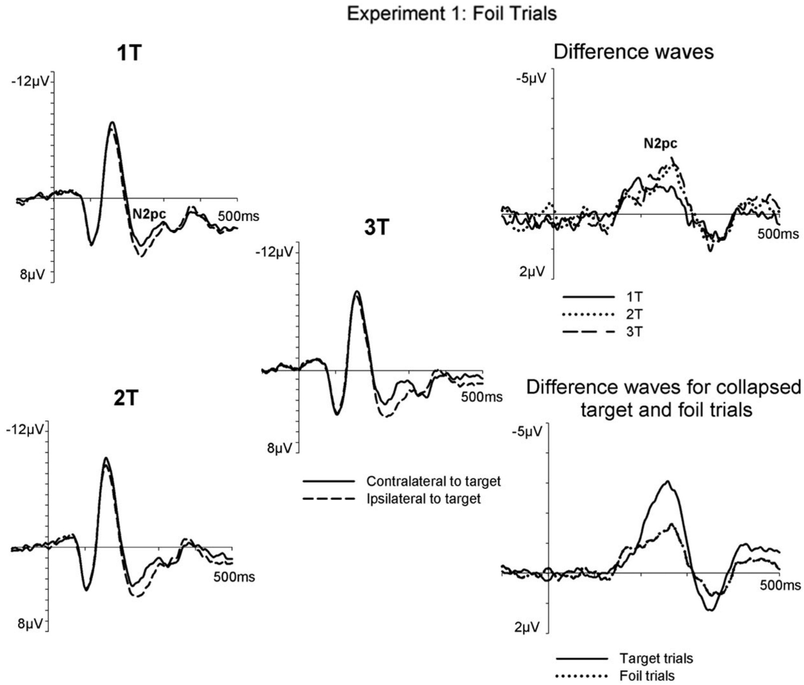

Figure 4. (Left and middle): Grand-average ERPs elicited in Experiment 1 in response to search arrays on foil trials at posterior electrodes PO7/8 contralateral and ipsilateral to a target item, shown separately for the 1T, 2T, and 3T condition. (Top right) N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERPs at PO7/8 for these three task conditions. (Bottom right) N2pc difference waveforms on target and foil trials, collapsed across the 1T, 2T, and 3T conditions.

contralateral and ipsilateral ERP mean amplitudes demonstrated reliable N2pc components to foils in all three task conditions: 1T: $t(15) = 5.75$, $p < .001$; 2T: $t(15) = 8.07$, $p < .001$; 3T: $t(15) = 6.29$, $p < .001$.

Figure 4 (bottom right panel), which includes N2pc difference waves to targets (solid line) and foils (dashed line), collapsed across the 1T, 2T, and 3T conditions, shows that foil-induced N2pc components were reduced in size relative to target N2pcs. To verify this, ERP mean amplitudes obtained in response to targets and foils in the 1T, 2T, and 3T conditions were analyzed together, with the factors task condition, laterality, and the critical new factor trial type (target vs. foil trials). A main effect of laterality, $F(1, 15) = 79.9$, $p < .001$, $\eta^2 = .84$, was accompanied by a highly significant interaction between trial type and laterality, $F(1, 15) = 17.4$, $p = .001$, $\eta^2 = .54$), thus confirming the attenuation of N2pc amplitude to foils versus targets during the 180–300-ms analysis window.

Early N2pc (180–200-ms poststimulus). During the initial ascending phase of the N2pc between 180 and 200 ms after search array onset, this component appears to be present not only for target trials, but also for foil trials. To assess this formally, ERP mean amplitudes obtained at PO7/8 for the 180–200-ms post-stimulus interval were analyzed separately for target and foil trials, for the factors task condition and laterality. For target trials, a main effect of laterality, $F(1, 15) = 16.3$, $p < .001$, $\eta^2 = .53$, confirmed that an N2pc was already reliably present during this early time window. There was no interaction between task condition and laterality, $F < 1$. Importantly, a reliable effect of laterality was also present for foil trials, $F(1, 15) = 24.2$, $p < .001$, $\eta^2 = .62$, demonstrating that foils also triggered an early N2pc during the 180–200-ms time interval. The interaction between task condition and laterality approached significance for foils, $F(1, 15) = 3.0$, $p = .065$, $\eta^2 = .17$. Paired *t* tests confirmed that early foil N2pc components were reliably present for all three task conditions: T1: $p < .001$; T2: $p = .018$; T3: $p = .033$. To further demonstrate that the early N2pc was triggered by all category-matching items, regardless of whether they were targets or foils, an additional analysis was conducted across both types of stimuli, with trial type (targets vs. foils) as an additional factor. There was a main effect of laterality, $F(1, 15) = 22.1$, $p < .001$, $\eta^2 = .60$, again confirming the presence of an early N2pc. Importantly, there was no interaction between trial type and laterality, $F(1, 15) = .07$, $p = .798$, indicating that this early N2pc component did not differ in size between target and foil trials.

Discussion of Experiment 1

Experiment 1 confirmed our previous observation (Wu et al., 2013) that item-based search is more efficient than category-based search. Targets were detected faster and N2pc components were larger during search for a single item (1T condition) than during search for any category-matching item (CT condition). However, the fact that reliable N2pc components were also triggered in response to category-defined targets provides further evidence that category-based, top-down guidance of attention can produce spatially selective effects at early perceptual stages of visual processing.

An important finding of Experiment 1 was that there were no systematic performance or N2pc differences between multiple-item search (Conditions 2T and 3T) and purely category-based search (Condition CT; see Figure 3). One possible account for this pattern of results is that search for two or three specific letters or digits was controlled by parallel independent representations of each target item and that attentional guidance by these multiple target templates was equally efficient as purely category-guided search. However, the N2pc results obtained on foil trials suggest a different interpretation. Foil items triggered reliable N2pc components (see Figure 4), indicating that target category-matching stimuli that were not part of the currently relevant target set were able to attract attention. These results suggest that top-down attentional guidance was at least in part category-based in this experiment. With respect to the time course of item-based versus category-based guidance, the pattern of results obtained during the early N2pc time window was particularly informative: Both targets and foils triggered reliable early N2pc components. The presence of an early N2pc to foils demonstrates that category-based guidance was remarkably fast and produced spatially selective effects in extrastriate visual cortex within less than 200 ms after search array onset. It is also important to note that early N2pc amplitudes did not differ between target and foil trials (see Figure 4, bottom panel). The absence of an interaction between trial type and laterality during the early N2pc time window suggests that the earliest stage of attentional target selection was controlled entirely by a category-defined task set. Item-based attentional control emerged slightly later: From about 200 ms after stimulus onset, N2pc components were reliably larger on target compared with foil trials, indicating that the match between an item in a specific search array and the currently active task set affected attentional processing at this stage.

The results of this experiment suggest that both item-specific and category-based, top-down control mechanisms are involved in the attentional selection of category-defined targets (letters among digits, or vice versa) and that category-based effects precede item-specific effects. The observation that attention was initially attracted by all items that matched the current target category, regardless of whether they were targets or foils can also account for the finding that target RTs did not differ between the 2T, 3T, and CT conditions. Further support for this conclusion comes from the observation that target-absent RTs were delayed by the presence of a foil item in the display, demonstrating that it was more difficult to reject a search array as a nontarget when it contained an item that matched the target category. It is notable that delayed target-absent responses on foil versus distractor-only trials were not only found in 2T and 3T blocks, but also during search for a single target item (Condition 1T). Interestingly, reliable early foil N2pcs were also present in this condition. Although target selection in 1T blocks could have been exclusively based on a physical match with an item-specific attentional template, delayed target-absent RTs and the presence of an N2pc on foil trials suggest that category-based attentional control processes were still active. There was however a marginally significant tendency toward a smaller foil N2pc during single-item search relative to multiple-item search (see Figure 4), which might suggest that the effects of category-based selection mechanisms are reduced when observers search for one specific target item. From approximately 200 ms after search array onset, target N2pcs were reliably larger during single-item search compared with multiple-item or category-based search (see Figure 3), demonstrating that target selection was most efficient when it could be guided by a physical match with one attentional template (see also Wu et al., 2013, for similar results).

Overall, Experiment 1 has revealed new insights into the interplay between category-based and item-based attentional guidance of visual search for letters or digits. Under conditions where both strategies are available to guide attention toward target locations, both appear to be activated in parallel. To further isolate the distinct contribution of category-based, top-down control to attentional target selection in visual search, it would be informative to contrast the findings obtained in this experiment with a situation where observers have to search for one or multiple target items, when a category-guided selection strategy is no longer available. This situation was investigated in Experiment 2.

Experiment 2

In Experiment 2, participants were instructed to search for one, two, or three specific letter targets in different blocks (1T, 2T, and 3T conditions). In contrast to Experiment 1, letter targets were now no longer presented among digits, but instead among letter distractors. Under these conditions, target selection can no longer be guided by a category-defined task set and therefore had to be controlled exclusively by item-specific representations of the currently task-relevant letters. If the top-down guidance of visual search by attentional templates for specific target objects is strongly capacity-limited (Olivers et al., 2011), there should be systematic differences in the efficiency of target selection between all three task conditions, and these should be reflected in behavioral as well as in N2pc differences. In particular, and in contrast to Experiment 1, target selection should be less efficient when three different letters are potential targets (3T condition) relative to search for two possible target letters (2T condition), due to the costs associated with maintaining three rather than just two item-specific representations in working memory. Similar to Experiment 1, target selection should be more efficient during search for a single target letter (1T condition) than during search for two or three letters.

Methods

Participants. Fifteen paid volunteers participated in this experiment. Two participants were excluded due to excessive eye-movements (more than 60% of trials removed due to eye movement artifacts), and another because tasks were not delivered in the correct counterbalanced order. All remaining 12 participants (M = 30.3 years old, SD = 4.4, range: 24–38 years old, eight men) had normal or corrected vision.

Stimuli, design, and procedure. These were identical to Experiment 1, with the following exceptions. Targets and distractors were now always drawn from the same category (letters), and the size of the letter set was increased to 12 items (see Figure 1). Because all search arrays now contained four letters, there was no Category Target (CT) condition nor foil trials. The experiment consisted of 15 blocks, with five successive blocks in each of the 1T, 2T, and 3T conditions, and task order was counterbalanced across participants. Each block contained 32 target trials and 32 distractor-only trials where none of the designated target letters appeared in the search array. As in Experiment 1, target N2pc onset latencies were compared across the three task conditions, using a 50% peak amplitude criterion and corrected *t*-values (t_c; Miller et al., 1998).

Results

Behavioral results. Figure 5 shows RTs on correct trials (left panels) and accuracy (right panels) for the three different task conditions (1T, 2T, 3T). On target trials, there was a main effect of task condition on RTs, $F(2, 22) = 26.5$, $p < .001$, $\eta^2 = .706$. Target RTs were faster in 1T blocks compared with 2T blocks (500 vs. 596 ms; $p < .001$) and faster in 2T blocks compared with 3T blocks (596 vs. 649 ms; $p < .03$). On distractor-only trials, an analogous pattern emerged for target-absent RTs. There was a main effect of task condition, $F(2, 22) = 34.6$, $p < .001$, $\eta^2 = .759$, as target-absent responses were faster in 1T compared with 2T blocks (512 vs. 667 ms; $p < .001$) and faster in 2T compared with 3T blocks (667 vs. 728 ms, $p < .05$).

There was also a main effect of task condition for accuracy on target trials, $F(2, 22) = 9.1$, $p = .001$, $\eta^2 = .452$. Accuracy was higher in the 1T compared with the 2T condition ($p < .005$), but did not differ reliably between 2T and 3T blocks. On distractor-only trials, accuracy was close to ceiling (see Figure 5) and did not differ between the three task conditions, $F < 1$.

ERP results. Figure 6 shows ERPs triggered in the 500 ms after search array onset at electrodes PO7/8 in response to target letters in the 1T, 2T, and 3T conditions. N2pc amplitudes were largest in the 1T condition, reduced in the 2T condition, and smallest in the 3T condition. N2pc onset also appeared to be delayed in the 3T condition. A repeated-measures ANOVA on ERP mean amplitudes in the N2pc time window (180–300 ms poststimulus) for the factors task condition (1T, 2T, 3T) and laterality (electrode contralateral vs. ipsilateral to the target item) revealed a main effect of laterality, $F(1, 11) = 37.5$, $p < .001$, $\eta^2 = .811$, and, critically, an interaction between task condition and laterality, $F(2, 22) = 10.9$, $p < .001$, $\eta^2 = .497$. Pairwise comparisons of contralateral and ipsilateral ERP mean amplitudes demonstrated reliable target N2pc components in all three tasks: 1T: $t(11) = 6.01$, $p < .001$; 2T: $t(11) = 6.63$, $p < .001$; 3T: $t(11) = 4.82$, $p = .001$. One-tailed *t* tests confirmed that N2pc amplitudes were larger in the 1T relative to the 2T condition, $t(11) = 2.29$, $p < .02$, and larger in the 2T relative to the 3T condition, $t(11) = 2.95$, $p < .006$. The N2pc emerged reliably earlier in 1T compared with 3T blocks (205 vs. 223 ms after search array onset: t_c (11) = 2.71, $p = .03$. The N2pc onset latency difference between 2T and 3T blocks (211 vs. 223 ms) was not reliable.

Discussion of Experiment 2

The findings of Experiment 2 were clear-cut: Search was most efficient when one specific letter served as target (1T condition), was reduced in efficiency when participants searched for two targets (2T condition) and was least efficient during search for three possible target letters (3T condition). RTs were faster in the 1T compared with the 2T condition and faster in the 2T compared with the 3T condition, and this was the case for target-present responses as well as for target-absent responses on distractor-only trials (see Figure 5). An analogous pattern was found for N2pc amplitudes, which were largest during search for one specific target letter and most strongly attenuated during search for three possible target letters (see Figure 6). Furthermore, N2pc onset was

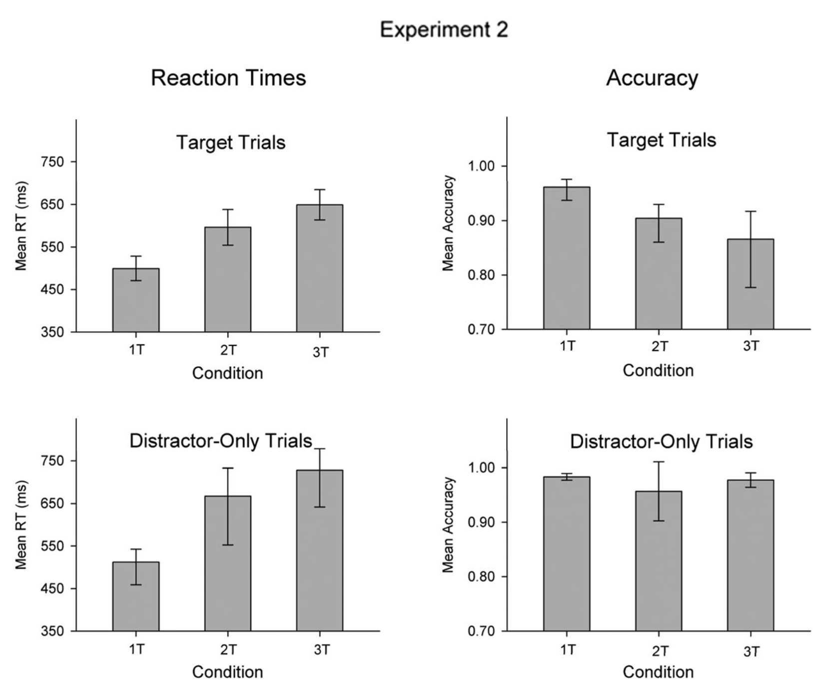

Figure 5. Mean RTs on correct trials (left panels) and mean accuracy (right panels) on target and distractor-only trials in Experiment 2, shown separately for the three different task conditions. Error bars indicate ±2 *SE* of mean.

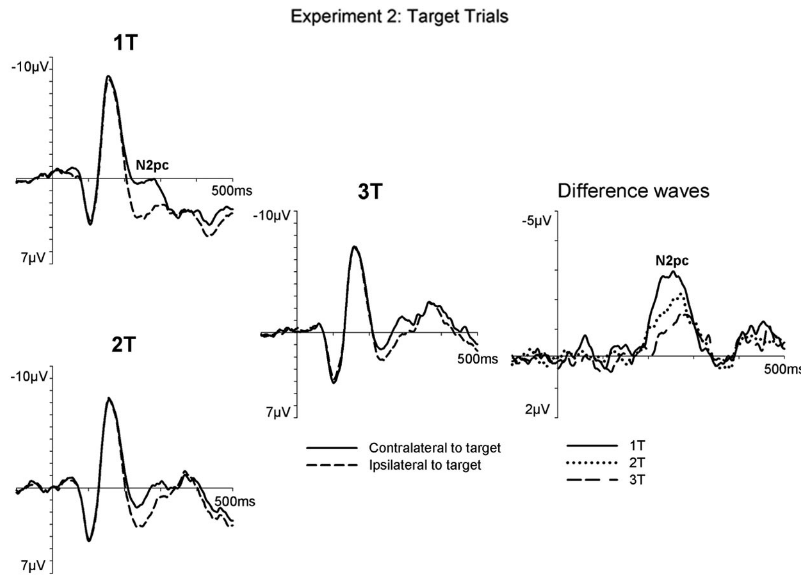

Figure 6. (Left and middle) Grand-average ERPs elicited in Experiment 2 in response to search arrays on target trials at posterior electrodes PO7/8 contralateral and ipsilateral to a target item, shown separately for the three task conditions. (Right) N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERP waveforms at PO7/8 for each task condition.

reliably delayed in the 3T compared with the 1T condition. These observations are in line with the hypothesis that attentional templates for specific target objects are strongly capacity-limited (Olivers et al., 2011): Target letters are selected less efficiently when two templates (rather than just one) have to be maintained in working memory, and search efficiency is further reduced when three target representations have to be kept active (see also Grubert & Eimer, 2012, for similar behavioral and N2pc results obtained during visual search for one vs. two target colors).

The main purpose of Experiment 2 was to serve as a contrast to Experiment 1, where both category-based and item-based strategies were available to guide attentional target selection. Eliminating category-based guidance in Experiment 2 produced a qualitatively different pattern of results. In Experiment 1, target RTs and N2pc components did not differ between the 2T, 3T, and CT conditions, and this was attributed to a common role of category-based guidance in all these conditions. In Experiment 2, RTs were reliably delayed and the target N2pc component was attenuated in the 3T compared with the 2T task. Furthermore, the N2pc was delayed in the 3T condition relative to the 1T condition in Experiment 2 but not in Experiment 1, which again is likely to reflect the absence versus presence of category-based attentional selection. Overall, these differences between Experiment 2, where target selection could only be guided in a strictly item-based fashion, and Experiment 1, where category-based guidance was available, provide additional strong evidence that category-defined top-down task sets did indeed play an important role for attentional target selection in Experiment 1.

General Discussion

The aim of the present study was to investigate the roles of item-based and category-defined task sets during search for one or multiple target objects. In Experiment 1, participants had to report the presence versus absence of letter targets that were presented among three digit distractors, or vice versa. In different blocks, they searched for one, two, or three specific items or for all items that matched the target category. The main finding of Experiment 1 was that alphanumerical category played a major role during attentional target selection not just during category-based search, but also during search for specific target items. When one, two, or three letters or digits were defined as targets, foil objects that matched the currently relevant target category but not the specific target items within this category triggered reliable N2pc components, indicating that they attracted attention even though they were task-irrelevant. Importantly, the N2pc was already reliably present during an early analysis interval (180–200 ms after search array onset) not just on target trials, but also in response to foils. If item-specific attentional control had preceded category-based control, an early N2pc should have been observed only for targets but not for foils. The presence of a significant early N2pc on foil trials thus demonstrates that category-based attentional guidance can be remarkably fast. In fact, the early phase of the N2pc did not differ between target and foil trials. This suggests that the initial allocation of attention was solely guided by category-defined top-down task sets and not by item-specific attentional templates, although it has to be acknowledged that this conclusion is based partially on a null effect (the absence of an interaction between trial type and laterality for the early N2pc in Experiment 1).

N2pc differences between target and foil arrays emerged at around 200 ms after display onset, indicating that at this stage, attentional target selection was affected by the match or mismatch of a candidate target item with stored target representations. This item-specific phase of attentional target selection was most efficient when it could be guided by a direct physical match with an attentional template (i.e., during search for a single target object), as reflected by the fact that RTs were fastest and N2pc components largest in the 1T condition of Experiment 1. However, these effects of item-specific attentional selection appeared later than the category-based effects (reflected by a significant early N2pc on foil trials), suggesting that category-guided attentional control might actually precede item-selective control, even when participants searched for only one specific target item.

The fact that a category-based search strategy was available in Experiment 1 can account for the absence of any performance or N2pc differences between blocks where participants searched for two or three specific objects, or for any category-matching item. This is expected if the initial attentional selection of targets was guided by category in all these conditions. Experiment 2 revealed the true efficiency costs of increasing memory set size in letter search under conditions where category-based selection was no longer possible. When participants searched for one, two, or three specific letters among letter distractors, RTs increased and N2pc components decreased and emerged later as a function of memory set size. In particular, and in contrast to Experiment 1, there were performance costs and corresponding N2pc amplitude differences during search for three compared with two possible targets, reflecting the demands of having to maintain an additional target representation in working memory. These results support previous claims that attentional templates are strongly capacity-limited (e.g., Olivers et al., 2011), although they may not be entirely consistent with the view that such templates can hold exactly one target representation at any given time. The differential effects of memory set size on search efficiency observed in Experiments 1 and 2 that are linked to the presence versus absence of category-based guidance provide direct evidence for the importance of object categories in the control of visual search.

The prominence and early onset of category-based attentional guidance and the fact that category-defined task sets were active also during single-letter search where target detection could have been based exclusively on a physical match with a stored attentional template may seem surprising, as object categories are generally regarded as doubtful or poor candidates for the top-down control of attention (e.g., Wolfe & Horowitz, 2004). The question of whether alphanumeric category can be used efficiently to guide attentional target selection has been studied for decades, but remains controversial. It is a generally accepted fact that visual search performance is better when targets and distractors belong to different categories (e.g., letters among digits) than when they are drawn from the same category. Under some condition, this "category effect" in visual search (Jonides & Gleitman, 1972) can even result in flat search functions (Egeth, Jonides, & Wall, 1972; see also Duncan, 1980), suggesting that alphanumeric category can be extracted rapidly for all items in a search array. On the basis of such observations, Duncan (1980) proposed a "late selection" theory of visual attention, which assumes that all stimuli are identified and categorized at an early preattentive parallel processing stage, but requires access to a later capacity-limited system in order to become available to conscious awareness and perceptual report. However, others have questioned the hypothesis that category membership is processed at an early parallel stage of perceptual processing and have attributed the category effect in visual search to the presence of low-level feature differences between items that belong to different categories (e.g., White, 1977; Krueger, 1984). If items within the same category are more similar to each other than items from different categories, the category effect in visual search might simply be due to lower target-nontarget similarity (Duncan & Humphreys, 1989) in across-category search displays. The observation that the category effect remains present even when the physical similarity of letters and digits is fully matched (e.g., Dixon & Shedden, 1987) suggests that this effect is not primarily caused by such low-level visual feature differences (see also Smilek, Dixon, & Merikle, 2006). If items can be categorized rapidly and in parallel at an early preattentive processing stage, category-guided attentional target selection may indeed be a highly efficient strategy during search for alphanumerical targets.

The current results are fully in line with this hypothesis. The observation that category-based guidance affected attentional object selection within 180 ms after search array onset, and even preceded the spatially selective effects of item-based top-down control, suggests that information about the category of alphanumerical stimuli becomes available very rapidly and can be immediately used in the guidance of spatial attention. The presence of early category-based N2pc effects during search for one specific target letter (Condition 1T in Experiment 1) suggests that attentional guidance by object categories is no less effective and may even be available earlier than guidance by item-specific attentional templates. The question of whether rapid category-based attentional control is specific to highly overlearned items such as alphanumerical stimuli or can also be used for other types of (perhaps less familiar) visual objects will be an important topic for future research (see Wu et al., 2013). Behavioral evidence from studies that used line drawings of animals and artifacts (Levin, Takare, Miner, & Keil, 2001) suggests that search for categorically defined objects is surprisingly efficient, although the question remains whether this may be driven by low-level visual feature differences between categories.

Another important question concerns the possible neural basis for the rapid category-based attentional guidance effects that were observed in the present study. Investigations into the brain mechanisms involved in the representation of perceptual categories have focused on the inferior temporal cortex, which is assumed to be responsible for visual object recognition, and its links to prefrontal cortex, which supports categorization judgments (e.g., Freedman, Poggio, Riesenhuber, & Miller, 2001; see also Miller, Nieder, Freedman, & Wallis, 2003, for a review). Meaningful objects such as words are assumed to be represented in a region in the left middle fusiform gyrus (visual word form area; e.g., Cohen & Dehaene, 2004), and more anterior parts of the left fusiform gyrus might be specifically involved in the neural representation of single letters (James, James, Jobard, Wong, & Gauthier, 2005). These left-hemisphere ventral temporal regions could be responsible for the parallel analysis and encoding of alphanumerical identity. Occipitotemporal areas involved in object recognition are known to have direct links to ventrolateral prefrontal cortex (see Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013, for a recent review), which is a likely source for the task-set dependent categorization of display items as letters versus digits, and targets versus nontargets. In this scenario, the rapid attentional selection of category-defined targets, as indexed by the early onset N2pc component to targets and foils in Experiment 1, reflects the spatially selective modulation of processing in extrastriate visual areas that is triggered by recurrent feedback signals from category-discriminative prefrontal cortex, which are themselves generated on the basis of visual form information provided by left inferior temporal cortex.

In summary, the current findings demonstrate that category-based control mechanisms play a central role in the attentional selection of alphanumeric target stimuli. Top-down task sets for categorically defined targets can trigger spatially selective modulations of visual–perceptual processing within less than 200 ms after stimulus onset, which is at least as fast or possibly even faster than the selective effects of item-specific attentional templates. Object categories can be very efficient in guiding attention toward the location of candidate targets in visual search.

References

Alexander, R. G., & Zelinsky, G. J. (2011). Visual similarity effects in categorical search. *Journal of Vision, 11,* 1–15. doi:10.1167/11.8.9

Alexander, R. G., & Zelinsky, G. J. (2012). Effects of part-based similarity on visual search: The Frankenbear experiment. *Vision Research, 54,* 20–30. doi:10.1016/j.visres.2011.12.004

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. *NeuroImage, 22,* 466–476. doi:10.1016/j.neuroimage.2003.12.049

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222. doi:10.1146/annurev.ne.18.030195.001205

Dixon, P., & Shedden, J. M. (1987). Conceptual and physical differences in the category effect. *Perception, & Psychophysics, 42,* 457–464. doi:10.3758/BF03209752

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87,* 272–300. doi:10.1037/0033-295X.87.3.272

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458. doi:10.1037/0033-295X.96.3.433

Egeth, H., Jonides, J., & Wall, S. (1972). Parallel processing of multielement displays. *Cognitive Psychology, 3,* 674–698. doi:10.1016/0010-0285(72)90026-6

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography & Clinical Neurophysiology, 99,* 225–234. doi:10.1016/0013-4694(96)95711-9

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science, 291,* 312–316. doi:10.1126/science.291.5502.312

Grubert, A., & Eimer, M. (2012). Qualitative differences in the guidance of attention during single-colour and multiple-colour visual search: Behavioural and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. doi:10.1037/a0031046

Hopf, J.-M., Luck, S. J., Girelli, M., Hagner, T., Mangun, G. R., Scheich, H., & Heinze, H.-J. (2000). Neural sources of focused attention in visual search. *Cerebral Cortex, 10,* 1233–1241. doi:10.1093/cercor/10.12.1233

Houtkamp, R., & Roelfsema, P. R. (2009). Matching of visual input to only one item at any one time. *Psychological Research, 73,* 317–326. doi:10.1007/s00426-008-0157-3

James, K. H., James, T. W., Jobard, G., Wong, A. C., & Gauthier, I. (2005). Letter processing in the visual system: Different activation patterns for single letters and strings. *Cognitive, Affective, & Behavioral Neuroscience, 5,* 452–466. doi:10.3758/CABN.5.4.452

Jonides, J., & Gleitman, H. (1972). A conceptual category effect in visual search: O as letter or as digit. *Perception, & Psychophysics, 12,* 457–460. doi:10.3758/BF03210934

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences, 17,* 26–49. doi:10.1016/j.tics.2012.10.011

Krueger, L. E. (1984). The category effect in visual search depends on physical rather than conceptual differences. *Perception, & Psychophysics, 35,* 558–564. doi:10.3758/BF03205953

Levin, D. T., Takarae, Y., Miner, A. G., & Keil, F. (2001). Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception, & Psychophysics, 63,* 676–697. doi:10.3758/BF03194429

Luck, S. J., Girelli, M., McDermott, M. T., & Ford, M. A. (1997). Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. *Cognitive Psychology, 33,* 64–87. doi:10.1006/cogp.1997.0660

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology, 31,* 291–308. doi:10.1111/j.1469-8986.1994.tb02218.x

Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision, 9,* 1–13. doi:10.1167/9.11.8

Miller, E. K., Nieder, A., Freedman, D. J., & Wallis, J. D. (2003). Neural correlates of categories and concepts. *Current Opinion in Neurobiology, 13,* 198–203. doi:10.1016/S0959-4388(03)00037-0

Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology, 35,* 99–115. doi:10.1111/1469-8986.3510099

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences, 15,* 327–334.

Smilek, D., Dixon, M. J., & Merikle, P. M. (2006). Revisiting the category effect: The influence of meaning and search strategy on the efficiency of visual search. *Brain Research, 1080,* 73–90. doi:10.1016/j.brainres.2005.07.079

White, M. J. (1977). Identification and categorization in visual search. *Memory & Cognition, 5,* 648–657. doi:10.3758/BF03197411

Wolfe, J. M. (2012). Saved by a log how do humans perform hybrid visual and memory search? *Psychological Science, 23,* 698–703. doi:10.1177/0956797612443968

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5,* 495–501. doi:10.1038/nrn1411

Wu, R., Scerif, G., Aslin, R. N., Smith, T. J., Nako, R., & Eimer, M. (2013). Searching for something familiar or novel: Top-down attentional selection of specific items or object categories. *Journal of Cognitive Neuroscience,25,* 719–729.

Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically defined targets. *Vision Research, 49,* 2095–2103. doi:10.1016/j.visres.2009.05.017

Received February 21, 2013
Revision received April 24, 2013
Accepted May 1, 2013 ■