The time it takes to turn a memory into a template

Anna Wilschut
VU University Amsterdam, The Netherlands

Jan Theeuwes
VU University Amsterdam, The Netherlands

Christian N. L. Olivers
VU University Amsterdam, The Netherlands

Visual search is typically guided by goals that are set within working memory. By varying the stimulus onset asynchrony (SOA) between a visual stimulus describing the target and the search display containing that target, previous studies have estimated how long it takes to implement such an attentional set. Here we investigate how long it takes to turn a visual memory representation (rather than a percept of the stimulus) into an attentional set. We used a memory-based postcuing procedure in which observers first encoded two colors into memory. A subsequent spatial cue indicated which of the two defined the relevant target color, followed by the search task. Experiment 1, which employed RT-based measures with unlimited viewing time, showed search slopes for relevant and irrelevant sets that suggested near instant guidance. However, Experiment 2 demonstrated that RT measures can suffer from severe underestimation. With temporally limited (and masked) viewing, accuracy scores showed that maximal search guidance was reached at about 400 ms. The results suggest that a visual memory can be turned into an attentional set within less than half a second. The implications for previous findings are discussed.

Introduction

Much of our everyday visual behavior consists of searching for specific objects that are relevant to the task at hand. Such visual search is assumed to be guided on the basis of a representation of the search goal, referred to as the attentional set or template, and that is further thought to be maintained by the working memory (Desimone & Duncan, 1995; Duncan & Humphreys, 1989). A question that is crucial for understanding visual search and its dynamics is how fast an attentional set can be selected from working memory and then applied to optimally guide search.

The time course of implementing an attentional set in visual search has been investigated earlier with the use of cued search tasks (Knapp & Abrams, 2012; Meyers & Rhoades, 1978; Schmidt & Zelinsky, 2011; Vickery, King, & Jiang, 2005; Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004). In these tasks the search target identity is indicated by a centrally presented cue at the beginning of each trial. Then, at various stimulus onset asynchronies (SOAs), the cue is followed by the search display, which potentially contains the target, presented together with a number of irrelevant items. The observer's task is to find the target and respond according to its presence or some other predefined property. Using this procedure, Wolfe et al. (2004) measured the effectiveness of exact visual cues (a picture of the target) as compared to categorical cues (a verbal description of the target) for an extensive range of SOAs. The reaction times (RTs) in the cued conditions were compared to a baseline condition in which the target was always the same within one block of trials. The results showed that exact visual cues led to a rapid improvement of overall RTs, and that it took only about 200 ms for search to become as fast as in the baseline condition. In contrast, guidance by verbal and categorical cue types was slower and never reached the level of the exact visual cues (see also Theeuwes & Van der Burg, 2011).

Note though that overall RTs per se are not necessarily indicative of the attentional set effectively being in place, as RTs may reflect other processing stages than the guidance of attention towards the target, including postselection decision and response selection stages (e.g., Mortier, Theeuwes, & Starreveld, 2005). Instead, the slopes of the RT × display set size function may then be more indicative of the operation of a search-guiding attentional set. As the slopes represent the search time increment for each item added, they provide a direct measure of search efficiency, with little, if any, effect of response-related
processes. A less consistent pattern was seen for the RT slopes in the Wolfe et al. (2004) study. In some experiments slopes decreased for up to 200 ms, but in other experiments the pattern of slopes was essentially flat from 0 to 50 ms SOA onwards, suggesting near instant guidance from the moment the cue is presented. These results seemed to be further corroborated by Vickery et al. (2005) who used a largely similar procedure and found a benefit for exact visual cues already at the shortest of their SOAs at 200 ms, with no further modulations as a function of time. This led to the conclusion that the attentional set is fully in place at 200 ms at the very latest.

However, there may be a more general problem with using RT-based measures of guidance (whether overall RTs or slopes), as was also noted by Wolfe et al. (2004) themselves. Apart from the fact that overall RTs may reflect processes other than attentional guidance, there is the problem that observers may postpone their search until the attentional set is fully implemented. The latter would result in an SOA effect on overall RTs (suggesting some time is indeed needed to implement a set—although this is then ambiguous as other processing stages may be affected, too), in combination with the absence of an SOA effect on slopes (suggesting near instant selection). Of course, depending on the individual strategy, the data may reflect a mixture of waiting and partial guidance. In addition, one may even wonder whether standard search slopes in principle provide the best index of attentional guidance. When slopes are calculated across all items in the display, they do not specify whether observed changes are due to changes in guidance towards the target feature or due to the time it takes to process just any item in general (be it target or distractor). For example, arousal or preparation effects as potentially induced by the cue also tend to increase over time and may speed up visual scanning processes (including eye movements, or postselective decision processes occurring at each visited item), without necessarily becoming more selective—that is, more tuned towards the target properties (Los & Schut, 2008; Posner & Boies, 1971).

Recently, Schmidt and Zelinsky (2011) used a different measure of attentional selectivity, which may be less sensitive to the aforementioned issues. Instead of RTs, Schmidt and Zelinsky measured the proportion of first saccades that correctly landed on the target. They argued that a saccadic accuracy measure would be a more sensitive index of guidance, as it would not include factors such as decision criteria. In addition, and interestingly, a long preview time was given to fully encode the cue, and instead of SOA, the time between the cue offset and the search display onset (interstimulus interval, ISI) was varied. By presenting the cue for as long as $3 \text{s}$, the authors aimed to exclude the possibility that the guidance measure would be affected by the cue processing time. The results showed that with exact pictorial cues, the proportion of initial saccades to the targets now peaked later than what was suggested by earlier RT estimates, namely for ISIs between 300 and 600 ms. After that, performance declined again. Categorical cues were even slower in affecting guidance, with performance reaching optimum somewhere between 600 and 3000 ms. However, this effect was so weak that it was virtually negligible. The fact that the saccade-based estimate for the exact visual cues is longer than that provided by Wolfe and colleagues (2004) as well as by Vickery and colleagues (2005), lends support to the idea that RT-based measures may sometimes underestimate the time that is needed to implement an attentional set.

A further interesting aspect of the Schmidt and Zelinsky (2011) study is the finding that the performance function appeared to be time-locked to the cue offset. This suggests that observers wait until the cue is no longer visible before they commit its representation to memory and then start searching on the basis of that memory. Alternatively, when the representation is already in memory at the end of the cue, its offset may be the signal to start turning the memory into an attention-guiding representation. The present study directly investigates how long it takes to turn a memory into an attention-guiding template. Moreover, we examine which effects the type of measurement (RT or accuracy) has on estimating the time course of implementing the template.

Whereas previous findings suggest that observers can turn a visually presented cue into a template within a few hundred milliseconds, we ask the question whether observers can use a sheer visual memory as the basis for the attentional set, and what the dynamics of this process are. These dynamics may be quite different than for exact pictorial cues. As previous studies suggest that search guidance increases rapidly and strongly for pictorial but not for verbal cues, it may be that direct feed-forward activation from the stimulus underlies the found temporal guidance effects, with little role for visual working memory. For example, as suggested by Wolfe et al. (2004), the rapid, enhanced guidance effect of the exact picture cues may be augmented by lower-level sensory priming between the cue and the search target. Such priming is thought to occur automatically, as a result of implicit learning across the stimulus repetitions (Kristjánsson & Campana, 2010; Maljkovic & Nakayama, 1994). Therefore, search guidance may develop very differently when boosted by a real visual trace than when based on working memory representation per se.

We ran two experiments using a memory postcueing procedure, which is illustrated in Figure 1. A trial started with the presentation of two colors at pseudo-random positions around the fixation point. Both colors
had to be attended and, together with their positions, encoded into working memory as observers did not know beforehand which of them would indicate the target set. Observers were given 1.5 s to load the colors into memory. Then, 1 s after the colors had disappeared, a spatial postcue indicated which of the two colors would be relevant for the subsequent search task (as it indicated the target set) and which color not (as it indicated a set of only distractor items). Previous studies have found that postcueing yields similar attentional correlates as does precueing (Dell’Acqua, Sessa, Toffanin, Luria, & Jolicoeur, 2010; Nobre, et al., 2004; Ruff, Kristjánsson, & Driver, 2007), so it appears useful as a tool for creating an attentional set on the basis of a memory representation. Furthermore, electroencephalographic (EEG) studies have provided evidence that such cues are maintained in visual working memory, as indicated by the contralateral delay activity (CDA; Carlisle, Arita, Pardo, & Woodman, 2011; Vogel & Machizawa, 2004; Woodman & Arita, 2011) of the event-related potential. Here we were interested in how long it takes to turn the visual memory for the cued color into an attentional set that is useful for searching a subsequent display. To that end, we varied the SOA between the postcue and the search display between 50 and 800 ms. Since both target and distractor colors were present in the precue and committed to memory and we measured performance time-locked to the spatial postcue, any modulations should be then due to implementation of the attentional set based on the chosen memory representation, and not on residual priming from the activation of the sensory trace.

Furthermore, to see how different response strategies affect the estimates of the speed with which an attentional set can be implemented, the two experiments differed with regards to viewing time and the dependent variable. In Experiment 1, participants had unlimited viewing time when searching the displays, and RTs were measured. This corresponds to the most typical way of measuring visual search guidance in the literature. However, as explained earlier, it runs the risk of underestimating the time it takes to implement the attentional set, as the unlimited viewing time allows observers to postpone their search until the attentional
set has been completely implemented. In Experiment 2 we therefore limited the viewing time of the search display by allowing only a brief presentation in combination with a mask. Here we measured search accuracy as a function of SOA. Thus, this is the first study to compare how search performance develops over time as a function of response conditions, within one and the same task. If RT measures indeed run the risk of underestimating the time to set up a template, we should see such estimates to differ for the accuracy measures in Experiment 2.

Finally, in order to be able to demonstrate a clear guidance towards the target feature, we separated the search slope for items carrying the relevant (target) color from the slope for items carrying the distractor color by independently manipulating the set size of the cued (relevant) and uncued (irrelevant) items in the search display. This has benefits over simply assessing the slopes across all items in the display, as previous studies have shown that search can be restricted to a subset of items, thus demonstrating top-down selectivity (Egeth, Virzi, & Garbart, 1984; Kaptein, Theeuwes, & van der Heijden, 1995; Theeuwes, Kramer, & Atchley, 1998). Accordingly, the more successful the implementation of the attentional set for the relevant items, the larger the slopes for that particular subset, and the smaller the slopes for the irrelevant subset of items. The time that observers spend in searching among relevant and irrelevant colors is thus reflected in the difference between the slopes for the two subsets. These slopes were calculated for RTs in Experiment 1 and for error scores in Experiment 2.

**Experiment 1**

In Experiment 1 postcuing was used in combination with a visual search task to measure the time it takes from selecting a color from memory to optimally using that color to guide search. In this experiment RT was used as a dependent measure. The set sizes for items in the target-relevant color and in the irrelevant color were independently varied, in order to directly assess the bias towards the target color. With time, the more selective the search becomes, the larger the search slopes for the relevant color subset, and the smaller the slopes for the irrelevant subset (see Kaptein et al., 1995).

**Methods**

**Participants**

Fourteen students (five males, 18–28 years old, \( M = 22 \) years) participated in exchange for money (8 €/hour) or course credits. All gave their informed consent prior to the experiment and reported having normal or corrected-to-normal eyesight. The study was conducted in accordance with the Declaration of Helsinki.

**Stimuli and apparatus**

Participants were seated in a dimly lit cubicle, at approximately 70 cm from a 19-in. monitor (1024 × 768 pixels, refreshing at 120 Hz). Stimulus presentation and response collection was done by E-prime software (Psychology Software Tools, Pittsburgh, PA). The background was gray (CIE: \( x = 0.308, y = 0.338 \), luminance: 29.95 cd/m²). A central white (\( x = 0.279, y = 0.311, 110.9 \) cd/m²) circle with a 0.06° radius served as fixation. The pre cue display was always a red (\( x = 0.62, y = 0.345; 8.89 \) cd/m²) and a green (\( x = 0.375, y = 0.542; 7.642 \) cd/m²) disk of 0.31° radius each, drawn on the circumference of a centralized imaginary circle with a 0.62° radius. The position of one of the disks was first randomly determined in each trial, and the second disk was then positioned at the opposite location. The pre cue was a smaller black dot with 0.19° radius, presented at the location of the relevant color pre cue. The search display items were red and green (colors as for the pre cues) bars of four different orientations (0°, 45°, 90°, and 135°). Search items were drawn on a 4 × 4 grid that covered an area of 8.01° × 8.01°. The 16 search array positions were organized in four quadrants that were separated from each other by a 2.47° distance. The distance between the positions within the quadrants was 1.85°. The target was either a red vertical bar or a green horizontal bar, determined randomly on each trial. In order to force selection (and not just let observers search for the unique object), the other exemplar of the target set was also always present and was drawn in a different quadrant of the grid. The rest of grid positions were filled by distractor bars (the number of filled positions being dependent on the set sizes), whose color, orientation, and position were random, with the restriction that each quadrant contained always an equal number of exemplars of each color.

**Procedure**

Factors SOA (50, 200, 400, and 800 ms), relevant color set size (four and eight) and irrelevant color set size (four and eight) were varied within participants and randomly mixed within blocks. There were 40 trials for each of these 4 × 2 × 2 conditions. The trial (see Figure 1) began by simultaneous presentation of the fixation dot and the pre cue of two color disks for 1.5 s. The pre cue was then removed, and the fixation dot was presented alone for 1 s. The post cue dot was then shown for 50 ms at the location of the relevant...
color circle of the precue. This postcue indicated the target color with 100% validity. The search display followed the postcue at one of the four SOAs, and remained on until the participant had responded. The next trial began after a 1000 ms blank period. The task of the participant was to search for the target bar (which was either a red vertical or a green horizontal) and to indicate the quadrant in which it was shown (four alternative forced-choice response). The responses were given on a standard US keyboard, with the keys “f,” “j,” “v,” and “n” corresponding to the quadrants “upper left,” “upper right,” “lower left,” and “lower right,” respectively. Participants were instructed to respond as fast as possible while maintaining a high accuracy level. Feedback was given about the accuracy and speed. There were first 40 practice trials, followed by ten experimental blocks of 64 trials each. Only the results from the experimental blocks were analyzed. The experiment took about 70 min in total, including the self-paced breaks between the blocks.

Results and discussion

Figure 2 shows the RTs, including the average RT across different set sizes. For the analyses, only correct responses that were given between 200 and 4000 ms post search display onset were included (excluding incorrect responses [8.6%], or excessively slow or fast responses [0.2%]). First, to assess overall performance, a repeated measures ANOVA (with \( \alpha = 0.05 \) and Greenhouse–Geisser-corrected \( p \) values when necessary) on the average RT showed that there was a significant effect of SOA on RTs, \( F(1.4, 18.1) = 28.8, p < 0.001 \). Pairwise \( t \) tests confirmed what can be seen in Figure 2, namely that RTs decreased from 50 ms to 200 ms SOA, \( t(13) = 7.4, p < 0.001 \), and from 200 ms to 400 ms SOA, \( t(13) = 3.5, p < 0.01 \). After 400 ms, there was no further change in RTs (\( p = 0.39 \), ns).

Then, in order to assess the effect of SOA on specific selectivity for the target color, we calculated RT slopes independently for the relevant and irrelevant color sets. These slope values are shown in Figure 3, together with the difference scores (i.e., the subtraction of the slopes for the irrelevant set from the slopes for the relevant set). A repeated measures ANOVA with relevance and SOA as factors showed that relevance had a significant effect on slopes, \( F(1, 13) = 63.1, p < 0.001 \). Observers were searching selectively among the relevant items, while they largely ignored the irrelevant color. However, neither the effect of SOA, \( F(3, 39) = 0.63, p = 0.60 \),
nor the interaction between relevance and SOA, \(F(3, 39) = 0.07, p = 0.96\), were significant. This would suggest that the selectivity was not modulated over time.

Performance accuracy was overall high (\(M = 91\%\)) and improved as SOA increased, \(F(3, 39) = 10.03, p < 0.001\). Pairwise \(t\) tests between sequential SOAs did not reach significance when corrected for multiple comparisons (\(M = 88\%\) at 50 ms vs. \(M = 91\%\) at 200 ms: \(p = 0.04\); 200 ms vs. \(M = 93\%\) at 400 ms: \(p = 0.12\); 400 ms vs. \(M = 94\%\) at 800 ms: \(p = 0.22\)). Accuracies were thus fully in line with the RTs.

Even though the subset-specific slopes were used as the measure of selectivity here, we also calculated the slopes in the standard way, across all items in the display for trials in which the set size for relevant and irrelevant items was equal (either both four or both eight, resulting in total set sizes 8 and 16), as to provide a comparison with previous cueing studies. These slopes were larger than zero at each SOA (\(p < 0.001\)) but did not vary over time (\(p = 0.67\)).

The search slope data of Experiment 1 suggest that a memory-based postcue successfully guided search to the cued subset, and did this almost instantaneously, already at the 50 ms SOA. This is implied by the significantly higher RT slopes for the relevant subset as compared to the slopes for the irrelevant subset, and this effect being invariant over time. Such a rapid implementation of a target template seems surprising, as one would assume at least a minimum amount of processing in order to process the cue, retrieve the color, and then turn it into a template. However, the result is consistent with earlier failures to find a time-based modulation of search slopes (Vickery et al., 2005; Wolfe et al., 2004). What could explain this lack of temporal effects? The overall RTs are very suggestive here. They do show a modulation, as overall speed improves for SOAs up to 400 ms. Initially, the RT decrement is nearly perfectly inverse to SOA increment (135 ms decrement from 50 to 200 ms SOA), reminiscent of the classic \(-1\) slope in cognitive bottleneck studies (Pashler, 1994). This supports the idea that observers do need some time to fully implement the target template, and that search thus starts later, but is efficient once started (Wolfe et al., 2004). In other words, observers do not start searching until they are ready to do so. In the RT task used in Experiment 1 postponing the search was an entirely viable strategy for the participant, as the search display was visible for an unlimited duration, until the participant had responded. In Experiment 2 we therefore tested whether temporal modulations in selectivity could be found if the search task was severely time limited and performance accuracy was measured, instead of RTs.

### Experiment 2

Experiment 2 was a replication of Experiment 1 but now with performance accuracy as the dependent variable. To make the task sensitive to errors and to compel observers to start searching as soon as possible, the search display duration was limited to 150 ms and backward masked.

### Methods

#### Participants

Fourteen new students (seven males, 18–28 years old, \(M = 24\) years, self-reported normal or corrected-to-normal eyesight) gave their informed consent and participated. Four participants were replaced with four new participants as they performed below 50% accuracy.

#### Stimuli and procedure

Stimuli and apparatus were otherwise identical to Experiment 1 except that the search display was now masked by a dense clutter of red and green bars (similar to the search items) that covered the whole search array (8.01° × 8.01°) for 300 ms. The search display was present for 150 ms. Participants were instructed to respond as accurately as possible with no strain put on speed. Feedback was given only about accuracy. In the practice block the search display duration was first set to 1000 ms from which it was reduced stepwise given that an accuracy level of at least 70% was maintained (while during the experimental blocks the search display duration was fixed). There were five experimental blocks, with 128 trials in each.

### Results

Figure 4 shows performance accuracy as a function of SOA. Trials with too fast (<200 ms) or slow (>4000 ms) responses, in total 0.7% of all trials, were excluded. A repeated measures ANOVA confirmed that SOA had a significant effect on total accuracy, \(F(3, 39) = 79.3, p < 0.001\). Overall performance improved progressively over time up to the 400 ms SOA (50 ms vs. 200 ms: \(p < 0.001\); 200 ms vs. 400 ms: \(p < 0.01\), after which the improvement became less prominent (400 ms vs. 800 ms: \(p = 0.05\)).

An analysis of guidance towards the target color as a function of time was then performed in the same way for the errors here as for the RTs in Experiment 1. Figure 5 shows the error slopes independently for the relevant and irrelevant color set sizes at different SOAs,
together with the difference scores (formed by subtracting the irrelevant color slopes from the relevant color slopes). Relevance had again a significant overall effect on slopes, $F(1, 13) = 54.47, p < 0.001$, while the effect of SOA was not significant, $F(3, 39) = 1.4, p = 0.26$. Importantly, this time, there was a significant interaction between relevance and SOA, $F(3, 39) = 3.34, p < 0.05$. The relevant subset slopes were larger than the irrelevant subset slopes at all except the first SOA (50 ms: $p = 0.23$; 200 ms: $p < 0.001$; 400 ms: $p < 0.001$; 800 ms: $p < 0.01$), indicating that selectivity was absent early in time and then increased with SOA. Comparing the difference scores between the SOAs showed that the increment in selectivity became reliable when the SOA was increased to 400 ms [50 ms vs. 400 ms SOA: $t(13) = 3.1, p < 0.01$; 50 ms vs. 200 ms SOA: $t(13) = 2.0, p = 0.06$]. Between 400 and 800 ms the difference score appeared to diminish again, but this effect was not significant [400 ms vs. 800 ms: $t(13) = 1.4, p = 0.2$].

Again, also the slopes for the total set size were calculated (similar to Experiment 1). These showed a different pattern than the subset slopes. While being overall significantly larger than zero ($p < 0.001$ at each SOA), no effect of time was found ($p = 0.26$). This demonstrates that slopes computed across all the items of a search array may not always be sufficiently sensitive to show temporal changes in guidance, as they do not differentiate between search among relevant and irrelevant items. Instead, one needs to specifically look at subsets carrying the target feature.

Similar to the RT results in Experiment 1, overall performance accuracy in Experiment 2 improved by postcuing the relevant color up to 400 ms SOA prior to search display onset, after which it leveled off. This time, and in contrast to Experiment 1, selectivity (as measured by the relevant set slope minus the irrelevant set slope) was also found to change over time. There was no selectivity for the target color at the shortest SOA of 50 ms, but it then increased as the SOA was prolonged, so that the maximal difference between the relevant and irrelevant subset search was reached at 400 ms.

General discussion

Two experiments were conducted to investigate the time that it takes to implement an attentional set from memory. For this purpose a memory postcuing procedure was combined with a visual search task at various SOAs. Whereas in previous studies when observers were given a single visual cue, which could be directly turned into a search template, our observers...
first had to load the visual information into memory before they got to know what to search for. This allowed us to measure, for the first time, how long it takes to turn a memory into a template for search. This is theoretically important because many theories of visual search assume a role for working memory in maintaining the template for visual search (Bundesen, 1990; Desimone & Duncan, 1995; Duncan & Humphreys, 1989; Wolfe, 1994). However, they do not specify whether just any active memory automatically serves as a template, or whether a working memory needs to undergo a transformation in order to successfully guide attention (e.g., Carlisle et al., 2011; Chun, 2011; Olivers & Eimer, 2011; Postle, 2006; Woodman, Luck, & Schall, 2007). Our experiments show the latter: First there are two memories active, after which one of them is turned into a search template. The findings suggest that this takes around 400 ms.

Furthermore, we are the first to show that RT measures, as used by previous studies (Knapp & Abrams, 2012; Vickery et al., 2005; Wolfe et al., 2004), are not ideally suited for measuring the time course of implementing an attentional set. The current data indicate that the RT slope can be affected by the strategy to postpone search until the attentional set is at least partially implemented, leading to an underestimation of the time to implement a template.

Finally, we are the first to provide an estimate for the time it takes to implement an attentional set while fully controlling for automatic priming effects from the cue. Note that most previous studies did not control for this. Wolfe et al. (2004) checked for priming effects from the previous trial, but they did not control for priming from the cue itself. Moreover, such priming effects from the previous trial were close to reliable in almost all experiments of the Wolfe et al. study, showing that priming in this paradigm is far from hypothetical. Wolfe et al. argued that whether or not we should regard such priming effects as top-down may be a matter of semantics. However, although priming effects may indeed be quite “intelligent” (e.g., they are sensitive to overall task settings: Fecteau, 2007; Olivers & Meeter, 2006), we want to suggest that they are different from the flexible settings that observers can adopt trial by trial on the basis of a new task instruction. That is the specific form of short-term, task-driven (rather than history or context-driven) top-down attentional control that we have been trying to measure here. In the present study the memory display contained both the irrelevant and relevant color (which both had to be attended and stored), equalizing any priming effects across the subsets. However, increased selectivity was found for the cued color only. Furthermore, we measured the modulations in selectivity time-locked to a postcue, which in itself did not carry any target features. This makes the sensory priming explanation even less likely. Consistent with this, the study of Schmidt and Zelinsky (2011) included an experiment in which they tried to control for priming effects by backward masking the cue. The results showed that masking the cue did not remove the guidance enhancement effect, suggesting that it is not dependent on a persistent visual trace per se (even though this would not exclude other types of priming that carry across the mask). The present results are in line with this conclusion and provide an even stricter test by showing that selectivity can be enhanced by a postcue that reactivates a template presented already a second ago.

The present results are generally compatible with previous findings (Knapp & Abrams, 2012; Meyers & Rhoades, 1978; Schmidt & Zelinsky, 2011; Wolfe et al., 2004) by showing that search guidance is enhanced by cueing the search target. The optimal cueing latency matches with the study of Schmidt and Zelinsky (2011) that used saccadic accuracy as a measure of guidance. Since both their and our study used a form of accuracy measure under brief viewing conditions (i.e., either the first fixation or limited by masking), neither is likely to suffer from underestimating the time to implement the attentional set. Note that Schmidt and Zelinsky observed a similar time course as we did, but as a function of a visual cue offset, and concluded that the enhanced guidance effect was due to a process of consolidating the cue to visual working memory. This consolidation process was further suggested to be supported by the temporary trace of the cue features in iconic memory, a memory type that may last a few hundreds of milliseconds after the stimulus offset (Sperling, 1960). Importantly, the present results show that similar rapid guidance enhancement is not dependent on the visual trace of the target-matching cue, but can be initiated also by selecting a representation from working memory with the help of a postcue, as the postcue signals which memory to turn into a template. We suggest something similar may have occurred in the Schmidt and Zelinsky (2011) study, where the initial visual representation (which is on for 3 s) is only turned into a template once the cue offset signals that the visual search display is imminent. Of interest here is another study looking at the time course of implementing a template by Wolfe, Horowitz, Palmer, Michod, and Van Wert (2010). They found that reaching optimal guidance takes always some minimum time, even when the target is invariably repeated from trial to trial. Wolfe and colleagues (2010) presented observers with search displays containing 16 disks, four of which were drawn in a different color. After a variable delay the disks were filled with search items, and the target was always presented in one of the four differently colored disks. The color of the target-
containing disks could be either variable or remain the same for a block of trials. The RTs in these conditions were compared to a baseline speed that was obtained when only the four disks were visible. The results showed that even though the blocked condition led to more efficient search in terms of RTs, it still took at least 200 ms to reach the baseline level. The representation of the target presumably must have been consolidated into a rather stable representation in such repeated target blocks, and thus it appears that the additional time was needed not to retrieve this representation but to apply it to the search task. This would mean that during a search experiment, observers do not continuously hold a template active even if it is readily available, but only activate the template when there is a signal to start searching (such as the appearance of the search display or any prior cue). These results thus support the idea that the transformation of a mere target memory into an attentional set contributes to the time course of search guidance enhancement.

The current study assumes that the search colors were stored in visual working memory and upon the presentation of the postcue, one of these visual memory representations was transferred into a search template. Given the relatively long time it takes to reach the optimal guidance (about 400 ms), one may argue that instead of visual representation in working memory, participants may have recoded the visual information into a categorical representation. For example, according to this interpretation participants must have stored the memory array in a verbal code (e.g., on a given trial: “red was top left, and green was right bottom”), one of which is retained after the postcue. If such a recoding strategy had been used, the procedure of the current experiment would be similar to the word cue condition of Wolfe et al. (2004) that did not show the actual target pictorially but instead gave a verbal cue (e.g., “black vertical”). Even though it is possible that participants may have transferred the visual information into some verbal code here this is not likely for several reasons. First, Theeuwes and Van der Burg (2011) have shown that cueing with the actual color feature (e.g., a patch of red) is much more efficient in visual (pop-out) search than cuing with a verbal label (the word “red”). Therefore, if the visual representation is directly available, there is every reason to assume that participants use this more optimal code. Second, the recoding of the visual representation into a verbal memory code (e.g., “top left red, bottom right green”) is in the current experiment difficult as the location of the two color patches changed from trial to trial along an imaginary circle. Recoding would be effortful and error prone and search would significantly suffer. Third, there is convincing evidence by Woodman and colleagues (2007) that when a target template for search changes from trial to trial, it is stored in visual working memory. In other words, Woodman et al. argue that the search template remains in a visual representation when it is variable (as it was in our experiments). This is also indicated by the fact that such cues generate a posterior contralateral delay activity of the event-related potential, which is indicative of visual working memory maintenance (Carlisle et al., 2011; Vogel & Machizawa, 2004; Woodman & Arita, 2011). When the search template is stable from trial to trial it can be shifted to a long-term memory like representation (Woodman, Carlisle, & Reinhart, 2013). Fourth, Schmidt and Zelinsky (2011) who used saccadic accuracy as an index of guidance demonstrated that word cues were basically inefficient in guiding search. If anything, word cues only led to a minor improvement in search occurring as late as between 600 to 3000 ms. When they used visual cues, the pattern of guidance was very similar to that found in the current work, with guidance reaching optimal levels after about half a second. The studies that have arrived at faster guidance times used RTs and are thus likely to suffer from underestimation. Although we cannot fully exclude a categorical component, these results from previous studies make it highly unlikely that in our experiments, participants only based their search on categorical representations.

A potentially interesting aspect of the enhanced guidance effect is that it appears to be short-lived in some cases. Schmidt and Zelinsky (2011; see also Meyers & Rhoades, 1978) reported a peak in selectivity for first saccades between 300 and 600 ms since cue offset, after which the proportion of initial saccades to the target declined again. Wolfe et al. (2004) showed a similar decline in performance at later SOAs in a substantial number of their data sets, although it never quite reached significance. In the current Experiment 2, a similar, but again nonsignificant, tendency for selectivity to decline at the longest latency was observed. The reason for the more pronounced transience in previous studies could have been that it may have been caused (partially) by the decaying visual trace of the cue. For example, the iconic memory representation has been estimated to vanish within few hundreds of milliseconds after the stimulus offset (Sperling, 1960), possibly related to the temporal persistence of the cortical activity as evoked by the visual stimulus (Ferber, Humphrey, & Vilis, 2005; Ruff et al., 2007). If performance in the studies using an exact cue was driven by such a stimulus-based effect, this may have been more short-lived than an attentional set that is built in the absence of target preview. However, transient enhancement may also partially reflect the attentional response itself, as triggered by the cueing signal. Such a transient attentional enhancement have typically been found in spatial cueing tasks (e.g.,...
Cheal & Lyon, 1991; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989), and was recently suggested to be related to the selection of a target from irrelevant distractor items (Wilschut, Theeuwes, & Olivers, 2011). Similarly, the attentional response as evoked by a feature cue may be boosted only for a brief period of time during which search is efficiently guided towards the target, but after this temporary boost, distractors start to gain strength and selectivity decreases (Olivers, 2010). The weak but consistent transient pattern across the studies investigating feature cueing rather than spatial cueing suggests that similar mechanism may operate in the feature domain, but future studies would be needed to directly investigate this.

In conclusion, the current study shows that an attentional set can be implemented rapidly based on selection from memory. RT measurements may obscure the time needed to complete this process. The timing of memory-based guidance is similar to what has been found before for visually presented cues, while issues of target versus distractor selectivity and sensory priming were controlled for.

Keywords: visual search, attention, working memory, postcuing, time course

Acknowledgments

This work was supported by an NWO Vidi grant (452-06-007) to CNLO.

Commercial relationships: none.
Corresponding author: Anna Wilschut.
Email: a.m.wilschut@vu.nl.
Address: VU University Amsterdam, The Netherlands.

References


