# Time course of target recognition in visual search

*Andreas Kotowicz[1,2]\*, Ueli Rutishauser[1] and Christof Koch[1,3]*

[1] Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA
[2] Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich, Switzerland
[3] Division of Biology and Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA, USA

**\*Correspondence:**
*Andreas Kotowicz, Institute of Neuroinformatics, University of Zurich and ETH Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.*
*e-mail: kotowicz@ini.phys.ethz.ch*

Visual search is a ubiquitous task of great importance: it allows us to quickly find the objects that we are looking for. During active search for an object (target), eye movements are made to different parts of the scene. Fixation locations are chosen based on a combination of information about the target and the visual input. At the end of a successful search, the eyes typically fixate on the target. But does this imply that target identification occurs while looking at it? The duration of a typical fixation (~170 ms) and neuronal latencies of both the oculomotor system and the visual stream indicate that there might not be enough time to do so. Previous studies have suggested the following solution to this dilemma: the target is identified extrafoveally and this event will trigger a saccade towards the target location. However this has not been experimentally verified. Here we test the hypothesis that subjects recognize the target before they look at it using a search display of oriented colored bars. Using a gaze-contingent real-time technique, we prematurely stopped search shortly after subjects fixated the target. Afterwards, we asked subjects to identify the target location. We find that subjects can identify the target location even when fixating on the target for less than 10 ms. Longer fixations on the target do not increase detection performance but increase confidence. In contrast, subjects cannot perform this task if they are not allowed to move their eyes. Thus, information about the target during conjunction search for colored oriented bars can, in some circumstances, be acquired at least one fixation ahead of reaching the target. The final fixation serves to increase confidence rather then performance, illustrating a distinct role of the final fixation for the subjective judgment of confidence rather than accuracy.

**Keywords: eye movements, object recognition, psychophysics, top-down attention, visual search, confidence judgement**

## INTRODUCTION

When searching for a known target in a visual scene, eye movements are guided by a combination of retinal input and information about the target stored in working memory. Depending on the task, the same scene can evoke very different scan paths. During free viewing, the most salient locations are preferentially fixated (Parkhurst et al., 2002; Peters et al., 2005; Mannan et al., 2009). When looking for a particular target, however, this pattern changes: locations that share features with the target are preferentially fixated (Williams, 1966; Yarbus, 1967; Zohary and Hochstein, 1989; Wolfe, 1994; Findlay, 1997; Motter and Belky, 1998; Bichot and Schall, 1999; Hooge and Erkelens, 1999; Beutter et al., 2003; Najemnik and Geisler, 2005; Navalpakkam and Itti, 2005; Einhauser et al., 2006; Rajashekar et al., 2006; Ludwig et al., 2007; Rutishauser and Koch, 2007; Tavassoli et al., 2007). That is, stimuli are fixated because of their behavioral relevance rather than their saliency. The more difficult the search, the longer it takes and the more fixations are required (Binello et al., 1995; Williams et al., 1997; Zelinsky and Sheinberg, 1997; Scialfa and Joffe, 1998). Throughout search, two decisions need to be made: where to next move the eyes (planning) and detecting the target. Planning has been extensively studied (Motter and Belky, 1998; Caspi et al., 2004; Najemnik and Geisler, 2005; Rutishauser and Koch, 2007; Zelinsky, 2008). Where to saccade next is largely determined afresh at every fixation with little carry-over of information from the last fixation (Wolfe, 1994; Findlay et al., 2001),

although search strategies such as proceeding in a clockwise fashion pre-determine some of these decisions (Peterson et al., 2007). After some time has passed, enough evidence about the target (the goal of the search) has accumulated and the search concludes successfully. At what moment in time, relative to fixation onset, do subjects possess enough information about the target to localize it? While this process is likely a gradual accumulation (possibly across multiple saccades), subjects at some point make a decision to stop the search and proceed to give a response. We asked our subjects to identify the location of the target and found that subjects looked directly at the target before stopping the search. They did so both when they were instructed to look at the target as soon as they knew where it was as well as when they were free to interrupt search at any given time by a button press. Does this imply that subjects first fixated an item on the screen, then identified it as the target and thus stopped the search? That is, does identification only proceed after fixation, serially? Alternatively, subjects could have first identified the target away from fixation (possibly over the course of several fixations) and then performed a saccade to its location for further processing. That is, target identification might occur in parallel with determining where to look next, a type of "look-ahead" processing. While previous work (Rayner, 1978; Palmer et al., 2000; Engbert et al., 2002; Godijn and Theeuwes, 2003; Caspi et al., 2004; McDonald, 2006; Kliegl et al., 2007; Angele et al., 2008; Baldauf and Deubel, 2008) as well as latency arguments (see below) suggest

that gradual accumulation is likely, this has not been conclusively demonstrated experimentally for two-feature (color, orientation) conjunction search.

One fundamental constraint on the speed of target recognition is imposed by the time required for information to arrive at the appropriate areas of the brain. The human visual system can detect the presence or absence of complex objects within a very short time (Potter, 1976; Thorpe et al., 1996). Stimulus-specific responses measured with surface EEG take at least 150 ms to emerge (Thorpe et al., 1996). The frontal eye fields (FEF) are known to be crucial for initiating voluntary eye movements. In macaque monkeys, the earliest single-neuron responses in FEF emerge after 75 ms. These very early responses are, however, neither stimulus nor response selective (Schmolesky et al., 1998). On the motor side, it takes at least 140 ms to stop the execution of a pre-planned eye movement in humans and monkeys (stop signal reaction time; Hanes and Carpenter, 1999; Emeric et al., 2007). However, in our experiment, the average fixation duration is only 170 ± 70 ms. It is thus conceivable that this is not enough time to detect a target and stop the search before the next saccade is executed. Here, we test this hypothesis.

We use a novel gaze-contingent (Perry and Geisler, 2002; Geisler et al., 2006) experimental paradigm to terminate search with millisecond accuracy after the eyes first come close to the target. We show that subjects' accuracy to detect the target is high and does not depend on dwell time on the target, even for times as short as 10 ms after landing on the target. Supporting earlier arguments directly, we find that information about the target is acquired at least one fixation ahead. Further, we show that subjects nevertheless choose to fixate the target in order to increase subjective confidence.

## MATERIALS AND METHODS

Twenty four subjects were paid for participating in the experiment. All had normal or corrected-to-normal vision and none were aware of the purpose of the experiment. The experiments were approved by the Caltech Institutional Review Board, and all subjects gave written informed consent. All subjects were tested for red-green color deficiency using 24 color plates (Ishihara, 2004). One subject had to be excluded due to color blindness (not included in number above).

### TASKS – SEARCH ARRAY

We created the search arrays by placing 49 items on a $7 \times 7$ grid with 3.25º and 2.25º spacing in the x and y directions, respectively (**Figure 1**). Uniformly distributed position noise of ±1.00º and ±0.50º was added to each grid position (*x* and *y* directions respectively). We then rearranged the items so they would fit inside an imaginary $4 \times 3$ grid (4 columns and 3 rows; see **Figure S1** in Supplementary Material). This imaginary grid was used for deciding whether to report a trial as correct or incorrect to the subject. We used this grid instead of the original $7 \times 7$ grid to decrease the accuracy necessary for correct target localization. In addition we used this grid after the experiment to calculate the chance performance of localizing a target correctly. The resulting average distance to the closest neighbor was 2.13º, while the minimal and maximal distances were 2.10º and 2.38º respectively.

There were four different search item types (e.g. all combinations of red/green and horizontal/vertical). Three out of those four unique item types were present in a particular search array.
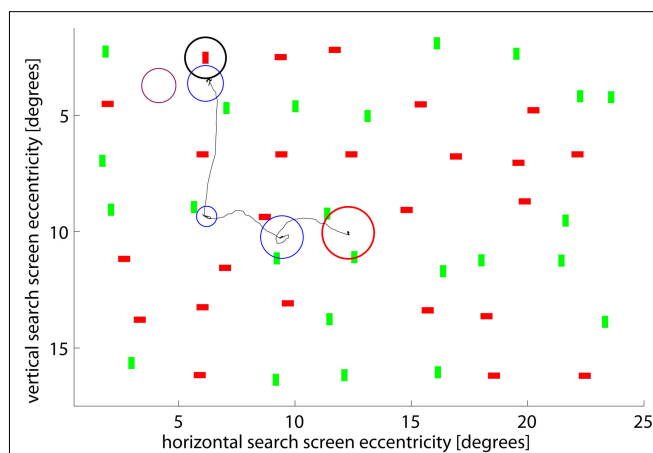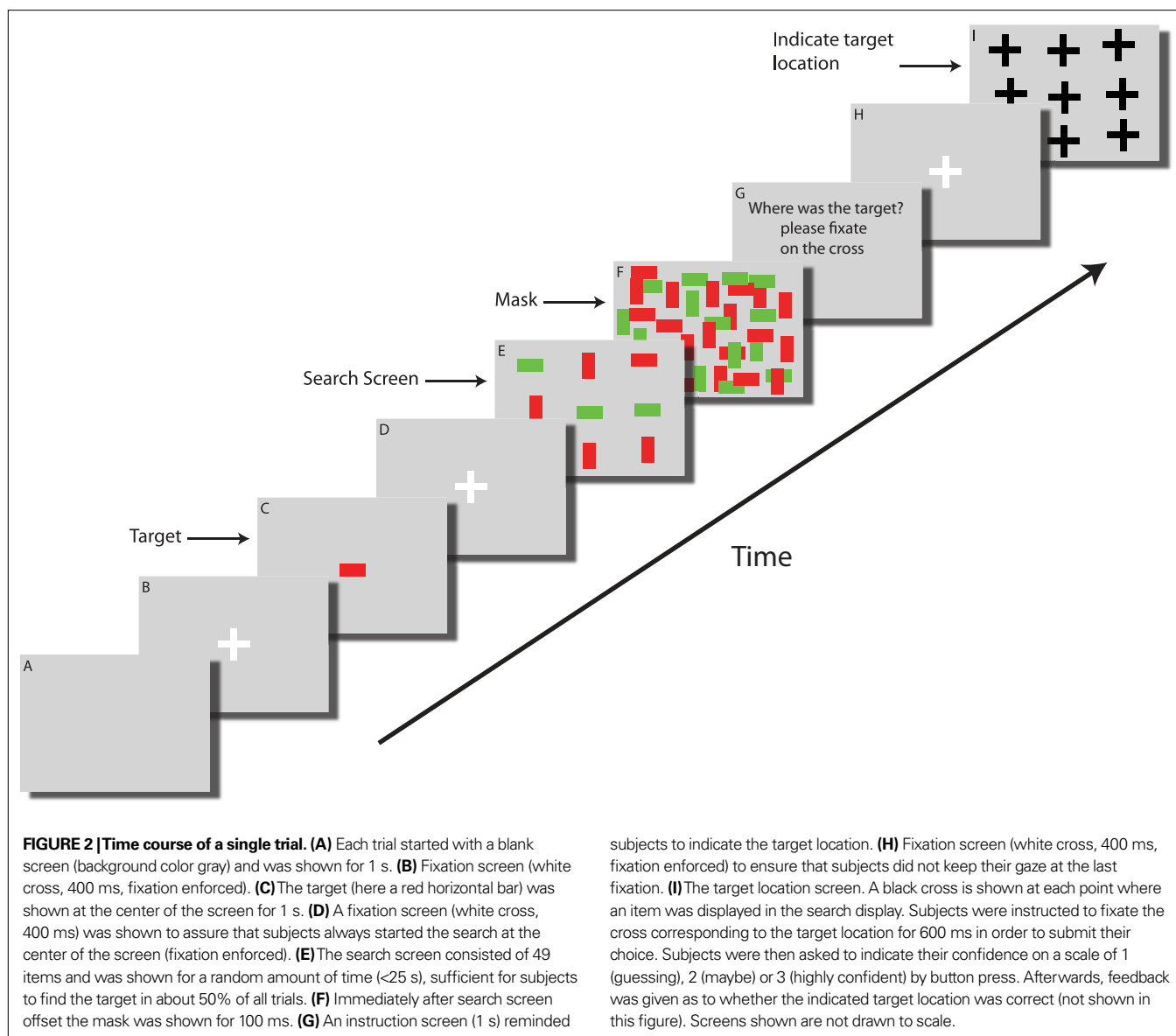


**FIGURE 1 | Example search screen and scan path of one subject.** The red circle marks the center of the screen, the black circle the location of the target (not shown to subjects). Blue circles show individual fixations – their radii are proportional to the fixation duration. A radius of 1.5° corresponds to 170 ms (shown in purple for comparison). The search was stopped after the subject fixated close to the target (3rd fixation). Screen eccentricity refers to the size of the objects, in units of visual angle, thus showing the size of the objects on the subjects' retina rather than on the screen itself.

The distractors were chosen such that half of them shared the first feature dimension with the target while the other half shared the second feature dimension (e.g. green/horizontal and red/vertical). Each search display consisted of 24 distractors and one target. The item size was 0.50º × 0.25º or vice versa.

### TASKS – PSYCHOPHYSICS

The screens shown to the subject are illustrated in **Figure 2**. At the beginning of each trial a blank screen was displayed for 1 s, followed by a white fixation cross at the center of the screen (400-ms display time). At the center of the next screen the target was presented for 1 s. To assure that subjects started the search at the center of the screen, we subsequently presented a second fixation cross which subjects had to fixate for 400 ms (within a 1.5º radius) to start the trial. If subjects failed to do so, recalibration was started automatically.

Depending on the experiment (see below), the search display (49 colored oriented bars) was present for a period of time between 20 ms and 25 s. Subjects were free to move their eyes (except during experiment 3, where subjects were required to maintain fixation within 1.5º of the center of the screen) and were instructed to find the target as quickly as possible. A trial was terminated either if subjects fixated within an area of 1.5º around the target for at least 400 ms (experiment 1 and 2) or if the trial timed out (whichever was first). The maximum time allowed for each trial was pre-determined before the start of the trial (range: 20 ms to 25 s, depending on the experiment; see below). If, during the maximal time allowed, subjects failed to identify the target by fixating on it for at least 400 ms, the trial was terminated regardless of the subjects' behavior. No manual interaction was required to terminate the trial except in the button press experiment.

**FIGURE 2 | Time course of a single trial. (A)** Each trial started with a blank screen (background color gray) and was shown for 1 s. **(B)** Fixation screen (white cross, 400 ms, fixation enforced). **(C)** The target (here a red horizontal bar) was shown at the center of the screen for 1 s. **(D)** A fixation screen (white cross, 400 ms) was shown to assure that subjects always started the search at the center of the screen (fixation enforced). **(E)** The search screen consisted of 49 items and was shown for a random amount of time (<25 s), sufficient for subjects to find the target in about 50% of all trials. **(F)** Immediately after search screen offset the mask was shown for 100 ms. **(G)** An instruction screen (1 s) reminded subjects to indicate the target location. **(H)** Fixation screen (white cross, 400 ms, fixation enforced) to ensure that subjects did not keep their gaze at the last fixation. **(I)** The target location screen. A black cross is shown at each point where an item was displayed in the search display. Subjects were instructed to fixate the cross corresponding to the target location for 600 ms in order to submit their choice. Subjects were then asked to indicate their confidence on a scale of 1 (guessing), 2 (maybe) or 3 (highly confident) by button press. Afterwards, feedback was given as to whether the indicated target location was correct (not shown in this figure). Screens shown are not drawn to scale.

The mask consisted of 800 randomly positioned red and green rectangles. It was shown for 100 ms to erase any retinal or iconic memory representation of the search display (Breitmeyer, 1984; Yantis and Jonides, 1996; Enns and Di Lollo, 2000). After the mask, an instruction screen was shown for 1 s followed by a fixation screen (white cross, 400 ms). Subjects were required to keep their gaze inside an area of 1.5° of the cross for the trial to continue. Trials where subjects failed to do so were discarded. This is to ensure that subjects did not keep their gaze at their last fixation (which might be the true location of the target). Afterwards, the search screen was shown again with all items replaced by black crosses. Subjects were asked to look at the location where they thought the target was. To indicate their choice, subjects needed to keep their gaze constant inside an area of 1.5° radius for 600 ms. The next screen asked subjects to indicate their level of confidence (confident, maybe, guessing) for this choice by button press. At the end, subjects received visual feedback (text, displayed for 500 ms) about whether their answer was correct or not.

A trial was considered correct (for the purpose of feedback) if the indicated location was within a 2.50° radius centered on the target. We used this coarse criterion to motivate subjects during the experiments but used a stricter rule for data analysis (see below).

Each subject performed only a single experiment type, which consisted of 8–10 blocks of 32 trials each. Prior to the experiment, subjects were given 20 practice trials (excluded from data analysis).

## TASKS – EXPERIMENTS

We performed three different experiments. Each subject only participated in one of the three. In experiment 1, subjects were instructed to find the target as quickly as possible and to indicate their choice by fixating on it. There were two categories of trials: normal and early terminated. In normal trials, subjects were allowed up to 25 s to find the target. In early termination trials, the search screen was removed before the subject was able to find the target. Early terminated trials would either terminate while the subject was fixating the target (early

termination on final fixation) or randomly through search (temporal early termination). We balanced the number of early terminated and normal trials so that in approximately 50% of all trials the duration was long enough for the subject to find the target (normal termination). We chose the next trial duration adaptively by sampling from a log-normal distribution that was generated by taking into account previous trial durations as well as their outcome (correct/incorrect). Subjects did not know whether a trial terminated because they fixated on the target (for 400 ms) or whether it was early terminated by the computer. In experiment 2 ("button press"), subjects were instructed to find the target as quickly as possible and to press a button as soon as they knew where it was. In experiment 2, the trial timeout was always 25 s. In both experiment 1 and 2, subjects were free to move their eyes during search. In experiment 3 ("fixation control"), the search screen was shown for a short amount of time (20–600 ms). Subjects had to maintain fixation within 1.5º of the center of the screen while searching for the target. In all three experiments, the same procedure followed after the end of a trial (mask, fixation, target location indication, confidence indication).

See Supplementary Material for written instructions given to subjects.

### EQUIPMENT – EYE TRACKING
Throughout all experiments we recorded subjects' right eye positions using a non-invasive infrared Eyelink-1000 (SR Research Ltd., Osgoode, ON, Canada) eye tracker with a sampling rate of 1,000 Hz. We used the manufacturer's software for calibration and validation (9-point calibration grid). The average radial resolution was 0.39º (the resolution of the eye tracker itself is 0.01º). Fixations were detected using the built-in fixation detection mechanism. We used the system's "real time" data acquisition mechanism which allowed us to react to eye movements with a delay of 2 ms. We confirmed this delay time by randomly sending timestamps during the experiments. **Figure 1** shows an example of a scan path.

### EQUIPMENT – SOFTWARE AND SCREEN
All experiments were implemented with Matlab (MathWorks, Natick, MA, USA) and the psychophysics toolbox version 3 (Brainard, 1997; Pelli, 1997) running on a Windows PC and a 19-in. CRT monitor (Dell Inc., Round Rock, TX, USA), which was located 80 cm in front of the subject. The maximal luminance (white screen) of the presentation screen was 29 cd/m$^2$. Maximal luminance for the green and red channels was 9.6 cd/m$^2$ and 6.1 cd/m$^2$ respectively. Ambient light levels were below 0.01 cd/m$^2$. The background of the screen was set to a light gray (14 cd/m$^2$) in order to reduce contrast. The display size was 25º × 20º. Subjects' heads were stabilized using a chin rest and a forehead rest to avoid head movement. The bit values used for red, green and gray were 255, 255 and 212 respectively. Red and green were not isoluminant. All experiments were run with a vertical screen refresh rate of 120 Hz; hence the refresh interval was roughly 8 ms.

### DATA ANALYSIS – INCLUDED TRIALS
We classified trials as correct if subjects indicated the correct target location (on the screen with crosses) with an accuracy of at least 1.5º. During the experiment, a trial was reported as correct to the subject if accuracy was at least 2.5º to not discourage subjects.

Results are reported (unless noted otherwise) as the mean over subjects and the standard error over the number of subjects. In case a subject contributed only 1 trial to the current data bin, this subject was not included into the analysis for this particular bin.

Six percent of all trials were excluded from analysis. These trials either timed out, were skipped because the target was too close to the center of the screen, because the subjects moved their eyes when they were not supposed to, or because eye movements were outside of the screen. We also excluded 38.2% of all fixation control trials (experiment 3) because we were only interested in trials where the search screen was shown for less than 400 ms.

The degree of difficulty for the search task was quantified by the number of fixations to find the target. All fixations between stimulus onset and mask onset were counted. We find that the average fixation duration, saccade duration and saccade size is quite stereotypic for all subjects (170 ms ± 70 ms, 44.3 ms ± 4.8 ms and 5.78º ± 0.85º, all ± SD). Consequently, the number of fixations to find the target is proportional to the time it takes to find the target.
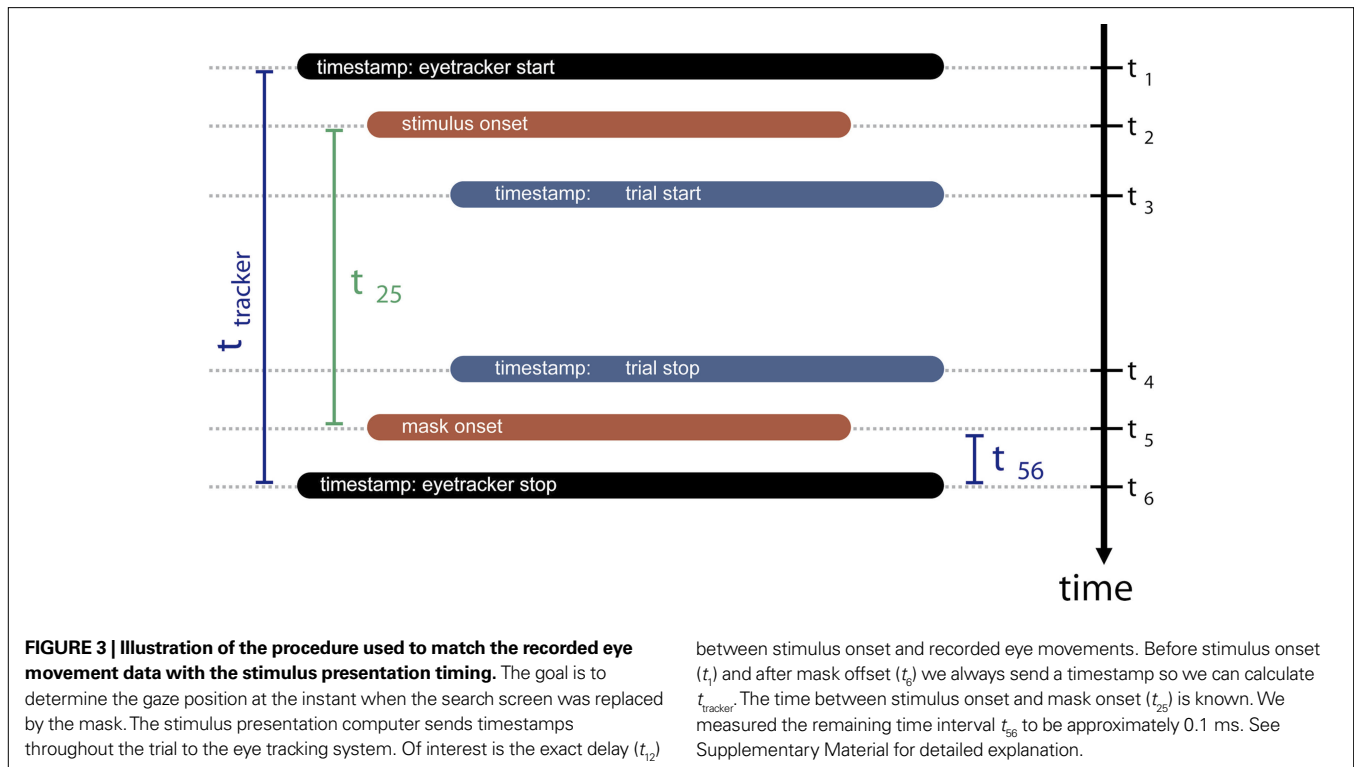
### EXACT TIMING OF STIMULUS ONSET AND EYE MOVEMENTS
We developed a method to match the actual stimuli presentation length (taking into account that a typical trial will terminate while the screen is in the middle of its refresh cycle ~4 ms) with the eye movements. The resulting uncertainty of this method is 1 ms. In order to not tamper accidentally with this stimulus/eye position matching, we always assume that the stimulus was present for this additional millisecond and thus include this additional data point into our analysis. Note that this is only done after the experiment is finished (offline).

During the actual experiment, some trials are terminated at random times before the subject identifies the target ("early terminated", see above). The procedure described here is used to retrospectively determine where the subjects' eyes were when the screen disappeared from the screen. This data is then used to identify where the subject's eyes were relative to the target and for how long. **Figure 3** illustrates the interaction between the eyetracker and the experimental computer during the course of a trial (see Supplementary Methods for details).

## RESULTS
### EXPERIMENT 1/LOOKING AT THE TARGET
We asked ten subjects to find a target among 48 colored oriented bars (**Figures 1 and 2**). The target was unique and shown to the subjects before each trial. Subjects were instructed to find the target as fast as possible and trials would terminate in one of the following three ways (experiment 1): (i) normal trials terminated after subjects looked at the target for 400 ms; (ii) trials terminated after subjects looked at target for less than 400 ms (early termination on final fixation; we tested 1–400 ms); (iii) trials terminated randomly throughout search (temporal early termination), which resulted in subjects' last point of view being 1.5º–20.0º away from the target. Technical constraints make it non-trivial to guarantee that stimuli changes triggered by eye movements are performed with millisecond accuracy. However, software we developed allowed us to do so with an effective time lag of 1 ms (see Materials and Methods for details). Following a mask and a central fixation cross (enforced) at the end of the trial, subjects were asked to look at

**FIGURE 3 | Illustration of the procedure used to match the recorded eye movement data with the stimulus presentation timing.** The goal is to determine the gaze position at the instant when the search screen was replaced by the mask. The stimulus presentation computer sends timestamps throughout the trial to the eye tracking system. Of interest is the exact delay ($t_{12}$) between stimulus onset and recorded eye movements. Before stimulus onset ($t_1$) and after mask offset ($t_6$) we always send a timestamp so we can calculate $t_{tracker}$. The time between stimulus onset and mask onset ($t_{25}$) is known. We measured the remaining time interval $t_{56}$ to be approximately 0.1 ms. See Supplementary Material for detailed explanation.

the location where they thought the target was, along with providing a three-level confidence rating. For normal trials, subjects needed on average $4.38 \pm 0.77$ fixations ($898 \pm 197$ ms) to find the target. The average saccade amplitude for correct trials was $5.56 \pm 0.85°$. Subjects correctly identified the target location for 83.3% of all normal trials, with chance corresponding to approximately 13.9%. The average confidence for all correct trials was 2.62 on a scale from 1 (guessing) to 3 (highly confident) (see Materials and Methods for details on how chance performance and confidence were calculated). We analyzed trials in which the search screen was removed and efficiently masked following a fixed interval (<400 ms) after the subjects acquired the target (that is, after they fixated the target within 1.5°; early termination on final fixation). The amplitude of the last saccade made to the target was on average $5.25 \pm 1.18°$ and was not significantly different from the average saccade amplitude (rank sum, $p = 0.43$). Trials with a last saccade amplitude of less than 4° were excluded to make sure that we do not look at effects caused by corrective saccades. Surprisingly (**Figure 4**), detection performance did not depend significantly on the delay between fixation onset and screen offset (ANOVA, $p = 0.61$, **Figure 4A**). This was also true for pair wise comparisons between the first and the 2nd–5th bin (t-test, $\alpha = 5\%$ uncorrected; **Figure 4A**). Thus, subjects always knew the target location with high accuracy. Even for trials in which the target was fixated only very shortly (<10 ms), performance was significantly different from chance ($p < 0.001$). This result also holds on a trial by trial basis (**Figure S2** in Supplementary Material). We calculated the mean confidence rating for all trials, regardless of whether the answer was correct or not. In contrast to performance, the mean confidence of all trials (correct and incorrect) did increase as a function of fixation duration from $2.3 \pm 0.2$ to $2.9 \pm 0.1$ (ANOVA, $p = 0.015$, **Figure 4B**). Mean confidence also increased if only correct trials were considered (ANOVA, $p = 0.028$; from $2.5 \pm 0.3$ to $3.0 \pm 0$, data not shown).

In case of randomly terminated trials (temporal early termination), performance and confidence did strongly depend on the distance between the last point of view and the target location (Performance and Confidence ANOVA, $p < 0.001$, **Figure 5A**). The median distance for correct trials was 4.1° and 10.1° for incorrect trials. A within-subject ANOVA revealed a significant difference between the two population of distances ($p < 0.001$, **Figure 5B**). Performance was not different from chance once the distance between target and last point of view was greater than 6° (**Figure 5A**).

So far, only the distance between the fixation point and the target was considered as a factor. Does performance also increase as a function of time? To evaluate this, we used the duration of the last fixation (before interruption). We expect a higher probability of being correct in trials where subjects fixated for longer at a distance where they could possibly detect the target (i.e. <6°). Indeed, we found that trials which terminated within a distance of 4–6° away from the target did show a dependency between the last fixation duration and correctness of the trial (within-subject ANOVA, $p = 0.013$, **Figure 5C**) and were independent of the previous saccade length (ANOVA, $p = 0.18$; data not shown). In contrast, trials which terminated within a distance of 1.5–4° away from the target, were independent of last fixation duration (ANOVA, $p = 0.97$; data not shown). Thus trials that did not terminate on the target were only correct if subjects fixated long enough during the preceding fixation, independent of the previous saccade size.

experiment 1 - early termination on final fixation
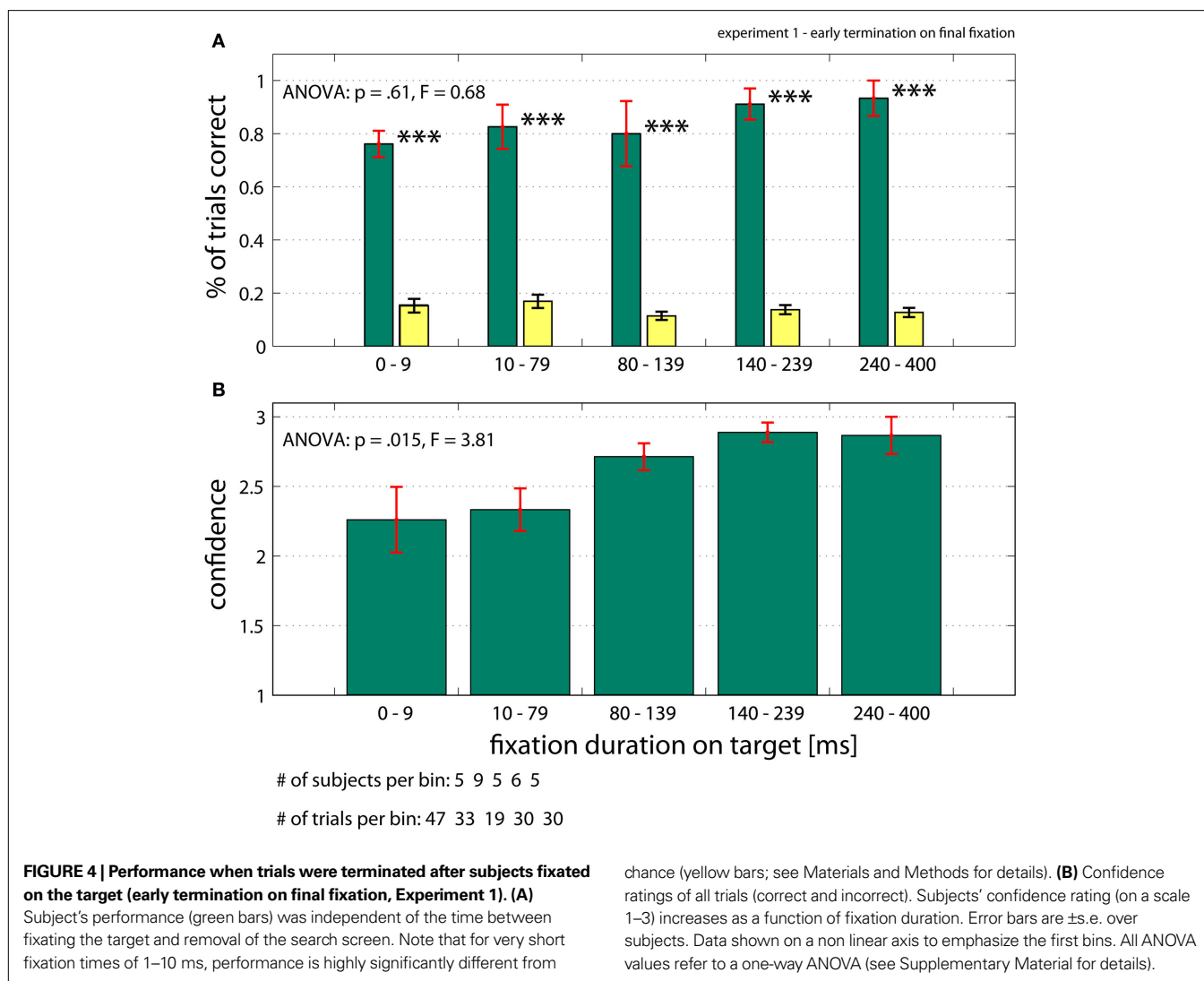
**FIGURE 4 | Performance when trials were terminated after subjects fixated on the target (early termination on final fixation, Experiment 1). (A)** Subject's performance (green bars) was independent of the time between fixating the target and removal of the search screen. Note that for very short fixation times of 1–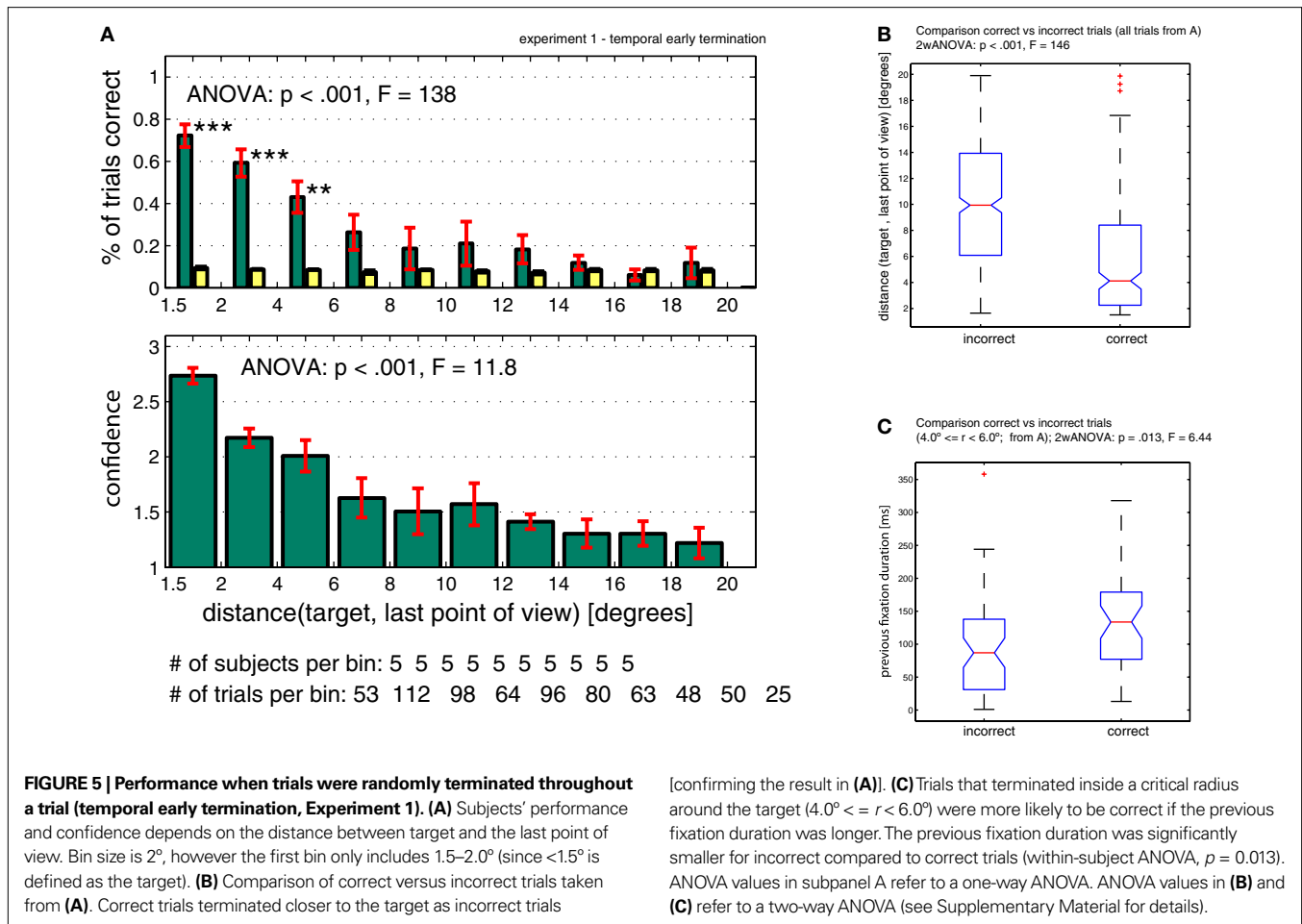10 ms, performance is highly significantly different from chance (yellow bars; see Materials and Methods for details). **(B)** Confidence ratings of all trials (correct and incorrect). Subjects' confidence rating (on a scale 1–3) increases as a function of fixation duration. Error bars are ±s.e. over subjects. Data shown on a non linear axis to emphasize the first bins. All ANOVA values refer to a one-way ANOVA (see Supplementary Material for details).

## EXPERIMENT 2/DETECTING THE TARGET

So far, subjects were required to look at the target to indicate successful localization of the target. Do subjects also look at the target in natural, unconstrained, search? We repeated the experiment with the instruction to localize the target and press a button as fast as possible, independent of whether or not subjects fixated the target and for how long (experiment 2). Afterwards, subjects (n = 3, who did not participate in the previous experiment) were asked to indicate the target position and confidence by pressing a button (see Materials and Methods). Subjects correctly identified the target location in 89.9% of all trials. In 91.2% of all correct trials, subjects fixated the target (within 1.5º) before pressing the button. The final fixation lasted on average 327 ms ± 96 ms, which is significantly longer than the average fixation duration (227 ± 142 ms). Thus, subjects looked at the target location for 100 ms longer than a typical search fixation, even though they were not required to do so (data not shown). 8.7% of all trials in experiment 2 terminated when subjects were not looking on the target (r > 1.5º). Still, subjects identified the target correctly in 80.8% of these cases.

## EXPERIMENT 3/DETECTION WITHOUT EYE-MOVEMENTS

How much time is required to identify the target in the absence of eye movements? To quantify this, we briefly flashed search arrays (for 20–600 ms) onto the screen and asked subjects (n = 11, who did not participate in the previous experiments) to identify the target's location (experiment 3). Subjects were required to maintain fixation within 1.5º of the center of the screen. The target was always located within 5.0º of the fixation point. If indeed the target is identified at the previous fixation, it should be impossible for subjects to locate the target as quickly as in the previous experiment (<10 ms). First, we performed this task with the mask immediately following the end of the trial as done previously (experiment 3 with mask, r < 5.0º). Subjects (n = 7) needed at least 375 ms to successfully localize the target (**Figure 6A**, 60% correct, p = 0.05 compared to chance). Performance as well as confidence increased as a function of time (ANOVA, p = 0.002 and p = < 0.001, respectively). Thus, the very short presentation times (<10 ms) sufficient to localize the target in active search (experiment 1, early termination on final fixation; successful identification in 76% of all trials) imply

**FIGURE 5 | Performance when trials were randomly terminated throughout a trial (temporal early termination, Experiment 1). (A)** Subjects' performance and confidence depends on the distance between target and the last point of view. Bin size is 2°, however the first bin only includes 1.5–2.0° (since <1.5° is defined as the target). **(B)** Comparison of correct versus incorrect trials taken from **(A)**. Correct trials terminated closer to the target as incorrect trials [confirming the result in **(A)**]. **(C)** Trials that terminated inside a critical radius around the target (4.0° < = $r$ < 6.0°) were more likely to be correct if the previous fixation duration was longer. The previous fixation duration was significantly smaller for incorrect compared to correct trials (within-subject ANOVA, $p = 0.013$). ANOVA values in subpanel A refer to a one-way ANOVA. ANOVA values in **(B)** and **(C)** refer to a two-way ANOVA (see Supplementary Material for details).

that the target location must have been acquired during previous fixations. Demonstrating this, subjects were at chance if we presented the exact same visual information for the same amount of time in the absence of previous fixations (**Figure 6A**). We repeated the experiment but this time targets were located outside the previously mentioned 5.0° around the center (experiment 3 with mask, $r > 5.0°$). Subjects were never able to identify the target better than expected by chance (12.2% ± 9.9%, **Figure 6B**) and performance only increased weakly with presentation time (ANOVA, $p = 0.02$). Subjects' confidence was never higher then guessing (1.49 ± 0.18) and did not significantly depend on time (ANOVA, $p = 0.13$), (data not shown). This result supports the previous finding that eye movements are necessary to perform the search task and that target detection is constrained to a specific detection radius.
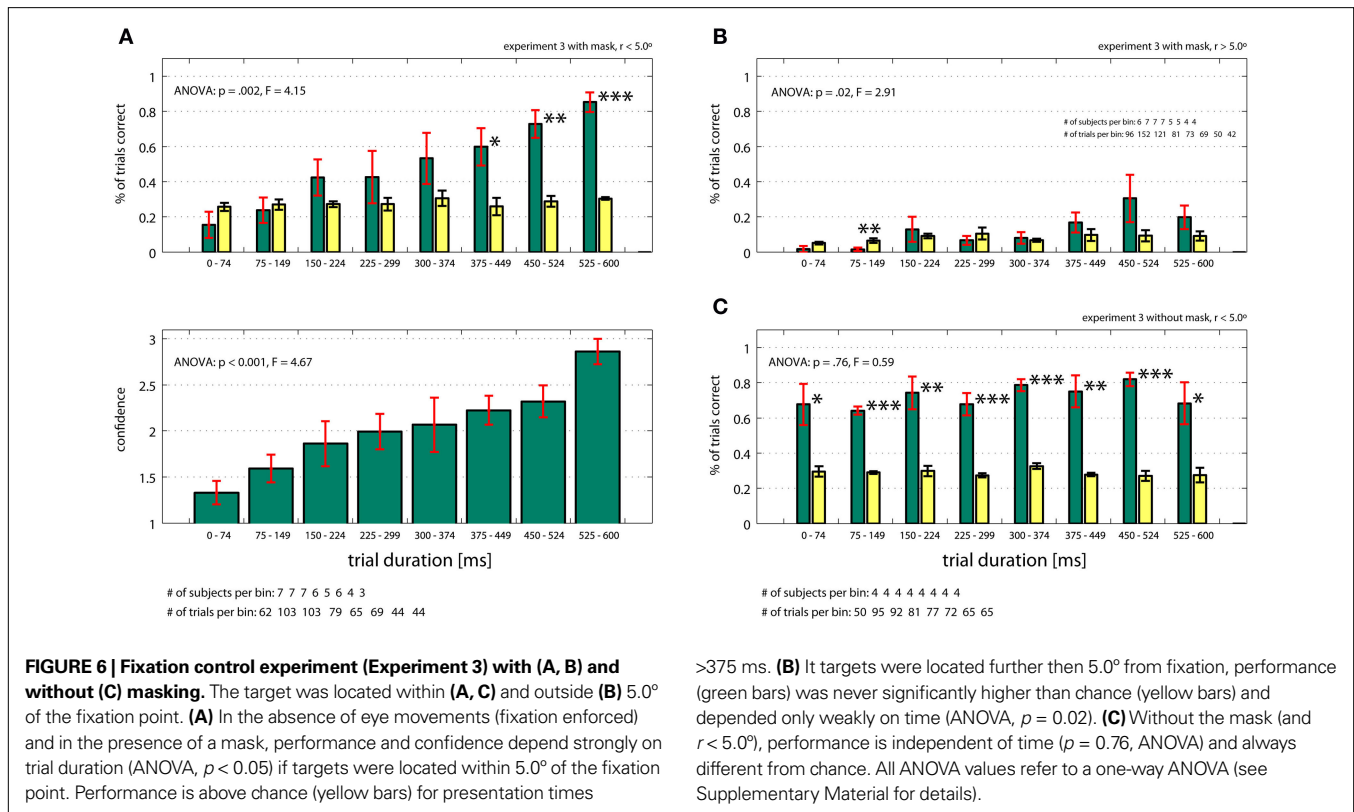
It could be argued that perhaps the mask we used is not effective in suppressing visual information. Thus, in a control experiment (4 subjects), we presented targets inside 5.0° of the fixation point, this time leaving out the mask at the end of the trial (experiment 3 without mask, $r < 5.0°$). It is known that target detection can proceed for very short presentation durations if no mask is presented (Thorpe et al., 1996; Fabre-Thorpe et al., 1998; Keysers et al., 2001; Li et al., 2002). It is thus expected that subjects will be able to locate the target in this condition. Confirming this, we found that subjects were able to identify the target with high confidence even at very short presentation times (**Figure 6C** 70% correct for <75 ms; $p = 0.05$ compared to chance; see **Figure S3** in Supplementary Material for trial by trial distribution). The masking effect is immediately apparent (compare **Figures 6A,C**). Therefore, we conclude that our mask was effective.

In our basic experiment, subjects reached performance levels between 76–93% correct (**Figure 4A**). Why were subjects never perfect? This gap in performance can, in part, be attributed to "return saccade" trials (which have been observed in monkeys as well as humans, i.e. (Motter and Belky, 1998; Peterson et al., 2001; Sheinberg and Logothetis, 2001). For the same task, we observed previously (Rutishauser and Koch, 2007) that subjects failed in approximately 12% of all trials to identify the target the first time they were directly looking at it. Subjects eventually returned to and identified the target ("return saccades"). Thus, landing on the target (without time constraint) does not guarantee detection of the target (also see below).

## DISCUSSION
Our results show that object detection in active conjunction search for a known target can occur with target fixation times under 10 ms (experiment 1, early termination on final fixation; **Figure 4A**): subjects correctly identified the target even if they fixated it for less

**FIGURE 6 | Fixation control experiment (Experiment 3) with (A, B) and without (C) masking.** The target was located within **(A, C)** and outside **(B)** 5.0° of the fixation point. **(A)** In the absence of eye movements (fixation enforced) and in the presence of a mask, performance and confidence depend strongly on trial duration (ANOVA, $p < 0.05$) if targets were located within 5.0° of the fixation point. Performance is above chance (yellow bars) for presentation times >375 ms. **(B)** It targets were located further then 5.0° from fixation, performance (green bars) was never significantly higher than chance (yellow bars) and depended only weakly on time (ANOVA, $p = 0.02$). **(C)** Without the mask (and $r < 5.0°$), performance is independent of time ($p = 0.76$, ANOVA) and always different from chance. All ANOVA values refer to a one-way ANOVA (see Supplementary Material for details).

than 10 ms and with a mask present at the end of a trial. This result stands in sharp contrast to our second finding that in the absence of eye movements, much longer times are required to successfully locate the target: the search screen had to be flashed on for at least 375 ms, otherwise subjects were unable to perform the task (experiment 3 with mask, $r < 5.0°$; **Figure 6A**). However, in the very same task, object detection can occur with very short display times if no mask is used (experiment 3 without mask, $r < 5.0°$; **Figure 6C**). Confirming previous results (Thorpe et al., 1996; Fabre-Thorpe et al., 1998; Keysers et al., 2001; Li et al., 2002; Serre et al., 2007), enough information can be acquired in these short (<75 ms) periods to detect the target.

How can these contradictory results be reconciled? These two cases differ in one substantial point: during active search (as opposed to at fixation), subjects previously fixated other positions in the display before landing on the target and therefore do have access to the relevant visual information, albeit at a non-zero retinal eccentricity, that is, away from the fovea, the point of highest acuity. This could be termed "look-ahead" processing. In contrast, there was not enough time to acquire sufficient information about the target while only fixating on it (as demonstrated by experiment 3 with mask, $r < 5.0°$). This interval during the previous fixation is indeed crucial as demonstrated by the temporal early termination experiment. Once subjects fixated at a critical distance away from the target, it is the previous fixation duration that determines the success of a trial (**Figure 5B**).

We therefore conclude that all the information necessary for the identification of our targets (here, comparing the color and orientation of an elongated bar with the target information in working memory) must have been acquired before the subject saccaded towards the target, at least one fixation ahead. This is why once the target is fixated, subjects are always able to correctly detect it even if the target was masked within 10 ms (by a mask that was disruptive enough to erase any retinal and iconic memory representation of the search display) (Breitmeyer, 1984; Kovacs et al., 1995; Yantis and Jonides, 1996; Enns and Di Lollo, 2000).

Note that our experiments show that subjects typically, if given the freedom to do so, fixate the target location before indicating knowledge of the target location (experiment 2). Thus, subjects fixate the target before they press the button to terminate the trial. However, our first and principal experiment (experiment 1) demonstrates that this is not necessary: if subjects are forced to identify the target location before ever fixating it (up to 6 deg away), they nevertheless succeed in identifying the target location (albeit with lower confidence). Thus, while the information necessary to identify the target was available, subjects did not terminate the search if they could choose to do so. One reason why this might be the case is for purposes of confidence, as shown below. Also note that in our first experiment, subjects were instructed to look at the target as fast as they could (to terminate the trial). Our experiment 2, however, shows that subjects also choose to look at the target if not forced to. In fact, they almost always (91% of cases) chose to look at the target before manually terminating the search.

## FIXATING THE TARGET INCREASES IDENTIFICATION CONFIDENCE
In our experiment, longer final fixation durations only increased the confidence, but not the performance, with which subjects made their decision (**Figure 4B**). Subjects accurately reported their confidence,

as demonstrated by a strong positive correlation between confidence and performance (see Supplementary Material). This finding led us to hypothesize that the reason why subjects looked at the target is to increase confidence. Otherwise there would be no reason for doing so, since the target identity is already known (at least by the saccadic system) before directly looking at it (within ~5º). We found supporting evidence for this hypothesis in the button press experiment (experiment 2). Subjects almost always (91% of trials) chose to look at the target before they pressed the button even though they were not instructed to do so (confirming previous results; Maioli et al., 2001). They also fixated on the target for longer (327 ms) before pressing the button than during the main experiment (experiment 1). This confirms that the instinctive behavior during active visual search is to first look at the target before confirming its location. However, in some cases subjects chose to identify the target without fixating on it. In these cases subjects were nevertheless highly accurate in identifying the target location (81% correct). Clearly peripheral identification is only possible if the items (given the resolution limits) can be discriminated when not directly fixating them. For targets that are much more difficult to discriminate than our colored bars (for instance bars that require the type of high spatial acuity information that's only accessible at the fovea), fixating the target directly might be necessary.

## INFORMATION PICK-UP FROM EXTRAFOVEAL LOCATIONS

It is thought that the neural substrates for controlling attentional and oculomotor shifts are largely the same (Rizzolatti et al., 1987; Corbetta et al., 1998). Furthermore, spatial attention shifts to the saccade target location prior to the onset of the saccade (Crawford and Muller, 1992; Hoffman and Subramaniam, 1995; Kowler et al., 1995; Deubel and Schneider, 1996). This, in turn, facilitates recognition processes at the location that is about to be fixated. Information necessary for recognition can thus be accessed away from the fovea if it is close enough to the current fixation (Geisler et al., 2006). Based on these findings and constraints imposed by response latencies (see below), it has been hypothesized that the target is typically recognized extrafoveally and not while fixating on it. Here we demonstrate experimentally that this hypothesis is true for the case of active visual search. Additionally we also demonstrate that while subjects detect the target before fixating on it, they nevertheless proceed to saccade to the target (in order to increase their confidence).

There are multiple tasks for which extrafoveal information acquisition has been shown, such as for saccade sequence programming or reading. In the former, eye movements are usually restricted to predefined locations between which subjects will saccade (preprogrammed saccade sequences) (Godijn and Theeuwes, 2003; Caspi et al., 2004; Baldauf and Deubel, 2008). This kind of task is different from natural visual search, where subjects freely explore the search space. In this respect, reading is more similar to visual search, where extrafoveal processing is a known phenomenon ("previewing") (Rayner, 1978; Engbert et al., 2002; McDonald, 2006; Kliegl et al., 2007; Angele et al., 2008). While it has repeatedly been hypothesized that similar processes are at work during recognition in visual search (and this assumption is implicitly built into many models of visual search), here we experimentally demonstrate this process. Note that saccade planning during search (where to fixate next) is also an extrafoveal process.

## GENERALIZATION TO OTHER TASKS AND COMPARISON TO PREVIOUS WORK

We used a conjunction search display consisting of 49 items, half of which shared a feature with the target. Based on this configuration, we found that targets could be identified successfully in the periphery up to 6 deg away. How do these results generalize to other search tasks? While we do not consider any other configurations in this paper, there are several factors that need to be considered. Clearly, the extent to which peripheral processing is possible depends on the target size, distractor density, discrimination difficulty and noise levels. If due to any reason the search becomes more difficult, we expect the effective radius within which this processing is possible to be reduced and at some point foveal processing will be required. On the other hand, if targets are sufficiently large and easy to discriminate, multiple fixations (active search) are less beneficial (Eckstein et al., 2001).

It has been shown previously that information about the target can be acquired peripherally (Viviani and Swensson, 1982; Scialfa et al., 1987; Scialfa and Joffe, 1998; Hooge and Erkelens, 1999; Eckstein et al., 2001; Caspi et al., 2004; Najemnik and Geisler, 2008; Zelinsky, 2008) and saccade target selection necessarily relies on parafoveal information. Our work adds to this literature by showing that this information is (for our task configuration) sufficient to make an accurate choice of target location, despite the fact that under typical circumstances subjects always look at the target. In contrast, it is not sufficient to make a high-confidence judgment, which requires foveal fixation. This is an important and novel distinction between the accuracy of the decision as such and its subjective confidence. Whether our findings are limited to the conjunction search display we used or generalize to other search configurations remains to be shown.

## MODELS OF VISUAL SEARCH

Our subjects found the target after an average of $4.38 \pm 0.77$ fixations. How many of the 49 items present in the display did the subjects likely consider during the search? In a purely random search without any attentional cues (and assuming that only one item is processed at every fixation), on average 24 saccades would be required to find the target. Clearly, subjects did not need to process every item since they knew which target they were looking for. Top-down attentional cues about the target reduce the number of fixations required to find the target (Williams, 1966; Findlay, 1997; Motter and Holsapple, 2000; Najemnik and Geisler, 2005). For similar search displays, subjects preferentially used color to guide the search (Williams, 1966; Motter and Belky, 1998; Williams and Reingold, 2001; Rutishauser and Koch, 2007). This reduces the number of possible items to 24 and thus the expected number of fixations to 12. (Rutishauser and Koch, 2007), used a quantitative model of search for the same display as used here to estimate that subjects process on average two to three items per fixation. This reduces the required number of fixations to 4–6. It is not known, however, whether these 2–3 items are processed sequentially or in parallel. Either way does not alter our finding and we thus remain agnostic about this issue. This theoretical consideration fits very well with the data we observed ($4.38 \pm 0.77$ fixations). Our previous work using the same stimuli, as well as similar experiments by others (Findlay et al., 2001), showed that all considered items are located inside a "search radius" of approximately 5–6° around the current fixation. The current experiments confirm this: whereas subjects are able to find the target inside a

5.0º circle around the fixation (**Figures 6A,C**, experiment 3 with and without mask, $r < 5.0º$), they are unable to do so for bigger radii (**Figure 6B**, experiment 3 with mask, $r > 5.0º$). Furthermore, during active search we found that performance does crucially depend on the distance between target and final point of view, in case trials are randomly terminated (**Figure 5A**) and performance is never found to be better then chance for distances bigger then 6.0º.

These results clearly emphasize the necessity of eye movements for this kind of search task. It is known that the density of items in a search array is a critical variable (Motter and Belky, 1998; Motter and Holsapple, 2000). To avoid this confound, we kept the density of items constant.

## NEURONAL CORRELATES

This "look ahead of fixation" processing of targets is compatible with a study reporting recordings from object-selective single neurons in the IT cortex in macaque monkeys (Sheinberg and Logothetis, 2001). Identical to our experiment, they considered fixations to be on target if they landed within 1.5º of the target (target acquisition). There were two key observations related to the time course of target identification: i) neurons exhibited a differential response to the target approximately 95 ms before the eyes landed on the target. This was only true if the monkey was about to fixate the target. ii) the same object selective IT neurons responded again once the eyes landed on the target.

Another fundamental constraint imposed on behavior is the onset latency of single neurons (Rolls and Tovee, 1994). It is frequently assumed that object-selective responses in IT cortex (such as the ones discussed above) are necessary for successful localization of complex objects. The minimum onset latency for monkey IT neurons is 85–95 ms (Nowak and Bullier, 1998; Naya et al., 2003) but can also be considerably later (Sheinberg and Logothetis, 2001). After the target is recognized, the search needs to be stopped (if fixating on the target). This requires approximately 140 ms (Hanes and Carpenter, 1999). Given these latencies and the average fixation duration, IT responses would be too late to induce an eye movement, since triggering a saccade requires the activity to reach other areas (such as FEF) first (see introduction). The different nature of these observations (macaque monkeys, different task) precludes any definitive conclusion in regards to our result. However, it is nevertheless of interest to note that these monkey IT single-neurons have properties which seem very similar to what we observed behaviorally. Also note that our task design did not allow express saccades

(which have latencies of as low as 120 ms; Kirchner and Thorpe, 2006). Of course, it remains speculative which specific neurons are activated during search in humans.

We used backward masking to terminate visual processing after a certain period of time. Measured psychophysically, the mask was effective in disrupting target recognition (see results). Neuronally, however, responses can occur long after mask onset and not all neuronal processing is disrupted (Thompson and Schall, 2000). Thus, the time of stimulus presentation is different from the duration of neuronal processing. Note that the latency argument (see above) does not require that neuronal processing is disrupted by the onset of the mask. Rather the argument states that, given the known response latencies of visual as well as motor neurons, the time available during a single fixation might not be sufficient for the entire process to complete (irrespective of whether the stimulus was masked or not).

## CONCLUSIONS

We conclude that, for our conjunction search task and configuration, the target identification process can be divided into two steps: a first round of processing, sufficient for target identification, takes places while the eyes fixate in the neighborhood but not on the target. A second round of processing starts once the final saccade arrives at its goal during which confidence increases. In cases where subjects fail to identify the target despite looking at it for ample time (>100 ms), either of these two processes might fail.

## SUPPLEMENTARY MATERIAL

The Supplementary Material contains a detailed descriptions of

i)   the statistical methods used for the data analysis;
ii)  the algorithm used for matching the eye movements and the stimulus presentation;
iii) the written experiment instructions as given to the subjects.

The Supplementary Material for this article can be found online at http://www.frontiersin.org/humanneuroscience/paper/10.3389/fnhum.2010.00031/

## REFERENCES

Angele, B., Slattery, T. J., Yang, J., Kliegl, R., and Rayner, K. (2008). Parafoveal processing in reading: Manipulating n+1 and n+2 previews simultaneously. *Vis. cogn.* 16, 697.

Baldauf, D., and Deubel, H. (2008). Properties of attentional selection during the preparation of sequential saccades. *Exp. Brain Res.* 184, 411–425.

Beutter, B. R., Eckstein, M. P., and Stone, L. S. (2003). Saccadic and perceptual performance in visual search tasks. I. Contrast detection and discrimination. *J. Opt. Soc. Am. A* 20, 1341–1355.

Bichot, N. P., and Schall, J. D. (1999). Effects of similarity and history on

neural mechanisms of visual selection. *Nat. Neurosci.* 2, 549–554.

Binello, A., Mannan, S., and Ruddock, K. H. (1995). The characteristics of eye movements made during visual search with multi-element stimuli. *Spat. Vis.* 9, 343–362.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.

Breitmeyer, B. G. (1984). *Visual Masking: An Integrative Approach.* New York: Oxford University Press.

Caspi, A., Beutter, B. R., and Eckstein, M. P. (2004). The time course of visual information accrual guiding eye movement decisions. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13086–13090.

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., Linenweber, M. R., Petersen, S. E., Raichle, M. E., and Van Essen, D. C. (1998). A common network of functional areas for attention and eye movements. *Neuron* 21, 761–773.

Crawford, T. J., and Muller, H. J. (1992). Spatial and temporal effects of spatial attention on human saccadic eye movements. *Vision Res.* 32, 293–304.

Deubel, H., and Schneider, W. X. (1996). Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vision Res.* 36, 1827–1837.

Eckstein, M., Beutter, B., and Stone, L. (2001). Quantifying the performance limits of human saccadic targeting during visual search. *Perception* 30, 1389.

Einhauser, W., Kruse, W., Hoffmann, K. P., and Konig, P. (2006). Differences of monkey and human overt attention under natural conditions. *Vision Res.* 46, 1194–1209.

Emeric, E. E., Brown, J. W., Boucher, L., Carpenter, R. H. S., Hanes, D. P., Harris, R., Logan, G. D., Mashru, R. N., Paré, M., and Pouget, P. (2007). Influence of history on saccade countermanding performance in humans and macaque monkeys. *Vision Res.* 47, 35–49.

Engbert, R., Longtin, A., and Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Res.* 42, 621–636.

Enns, J. T., and Di Lollo, V. (2000). What's new in visual masking? *Trends Cogn. Sci. (Regul. Ed.)* 4, 345–352.

Fabre-Thorpe, M., Richard, G., and Thorpe, S. J. (1998). Rapid categorization of natural images by rhesus monkeys. *Neuroreport* 9, 303–308.

Findlay, J. M. (1997). Saccade target selection during visual search. *Vision Res.* 37, 617–631.

Findlay, J. M., Brown, V., and Gilchrist, I. D. (2001). Saccade target selection in visual search: the effect of information from the previous fixation. *Vision Res.* 41, 87–95.

Geisler, W. S., Perry, J. S., and Najemnik, J. (2006). Visual search: the role of peripheral information measured using gaze-contingent displays. *J. Vis.* 6, 1.

Godijn, R., and Theeuwes, J. (2003). Parallel allocation of attention prior to the execution of saccade sequences. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 882–896.

Hanes, D. P., and Carpenter, R. H. (1999). Countermanding saccades in humans. *Vision Res.* 39, 2777–2791.

Hoffman, J. E., and Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Percept. Psychophys.* 57, 787–795.

Hooge, I., and Erkelens, C. (1999). Peripheral vision and oculomotor control during visual search. *Vision Res.* 39, 1567–1575.

Ishihara, S. (2004). *Ishihara's Tests for Colour Deficiency 24 Plates Edition.* Tokyo: Kanehara and Co. Ltd.

Keysers, C., Xiao, D. K., Foldiak, P., and Perrett, D. I. (2001). The speed of sight. *J. Cogn. Neurosci.* 13, 90–101.

Kirchner, H., and Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Res.* 46, 1762–1776.

Kliegl, R., Risse, S., and Laubrock, J. (2007). Preview benefit and para-foveal-on-foveal effects from word n+2. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1250.

Kovacs, G., Vogels. R., and Orban, G. (1995). Cortical correlate of pattern backward masking. *Proc. Natl. Acad. Sci. U. S. A.* 92, 5587.

Kowler, E., Anderson, E., Dosher, B., and Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Res.* 35, 1897–1916.

Li, F. F., VanRullen, R., Koch, C., and Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9596–9601.

Ludwig, C., Eckstein, M., and Beutter, B. (2007). Limited flexibility in the filter underlying saccadic targeting. *Vision Res.* 47, 280–288.

Maioli, C., Benaglio, I., Siri, S., Sosta, K., and Cappa, S. (2001). The integration of parallel and serial processing mechanisms in visual search: evidence from eye movement recording. *Eur. J. Neurosci.* 13, 364–372.

Mannan, S. K., Kennard, C., and Husain, M. (2009). The role of visual salience in directing eye movements in visual object agnosia. *Curr. Biol.* 19, 247–248.

McDonald, S. A. (2006). Parafoveal preview benefit in reading is only obtained from the saccade goal. *Vision Res.* 46, 4416–4424.

Motter, B. C., and Belky, E. J. (1998). The guidance of eye movements during active visual search. *Vision Res.* 38, 1805–1815.

Motter, B. C., and Holsapple, J. W. (2000). Cortical image density determines the probability of target discovery during active search. *Vision Res.* 40, 1311–1322.

Najemnik, J., and Geisler, W. (2008). Eye movement statistics in humans are consistent with an optimal search strategy. *J. Vis.* 8, 4.

Najemnik, J., and Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature* 434, 387–391.

Navalpakkam, V., and Itti, L. (2005). Modeling the influence of task on attention. *Vision Res.* 45, 205–231.

Naya, Y., Yoshida, M., and Miyashita, Y. (2003). Forward processing of long-term associative memory in monkey inferotemporal cortex. *J. Neurosci.* 23, 2861–2871.

Nowak, L. G., and Bullier, J. (1998). "The timing of information transfer in the visual system," in *Cerebral Cortex,* eds J. J. Kaas, and K. Rockland (New York: Plenum Press), 205–241.

Palmer, J., Verghese, P., and Pavel, M. (2000). The psychophysics of visual search. *Vision Res.* 40, 1227–1268.

Parkhurst, D., Law, K., and Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Res.* 42, 107–123.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442.

Perry, J. S., and Geisler, W. S. (2002). "Gaze-contingent real-time simulation of arbitrary visual fields," in Human Vision and Electronic Imaging: Proceedings of SPIE, San Jose, CA.

Peters, R. J., Iyer, A., Itti, L., and Koch, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Res.* 45, 2397–2416.

Peterson, M., Beck, M., Vomela, M., Petrie, H., Kornbrot, D., Phillips, J., Boles, D., Pieper, W., Ludwig, I., and Pollatsek, A. (2007). Visual search is guided by prospective and retrospective memory. *Percept. Psychophys.* 69, 123.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., and McCarley, J. S. (2001). Visual search has memory. *Psychol. Sci.* 12, 287–292.

Potter, M. (1976). Short-term conceptual memory for pictures. *J. Exp. Psychol. Hum. Learn. mem.* 2, 509.

Rajashekar, U., Bovik, A., and Cormack, L. (2006). Visual search in noise: revealing the influence of structural cues by gaze-contingent classification image analysis. *J. Vis.* 6, 7.

Rayner, K., (1978). Eye movements in reading and information processing. *Psychol. Bull.* 85, 618–660.

Rizzolatti, G., Riggio, L., Dascola, I., and Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia* 25, 31–40.

Rolls, E. T., and Tovee, M. J. (1994). Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proc. R. Soc. Lond. B* 257, 9–15.

Rutishauser, U., and Koch, C. (2007). Probabilistic modeling of eye movement data during conjunction search via feature-based attention. *J. Vis.* 7, 1–20.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., and Leventhal, A. G. (1998). Signal timing across the macaque visual system. *J. Neurophysiol.* 79, 3272–3278.

Scialfa, C., Kline, D., and Lyman, B. (1987). Age differences in target identification as a function of retinal location and noise level: Examination of the useful field of view. *Psychol. Aging* 2, 14–19.

Scialfa, C. T., and Joffe, K. M. (1998). Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Percep. Psychophys.* 60, 1067–1082.

Serre, T., Oliva, A., and Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6424–6429.

Sheinberg, D. L., and Logothetis, N. K. (2001). Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision. *J. Neurosci.* 21, 1340–1350.

Tavassoli, A., Van Der Linde, I., Bovik, A., and Cormack, L. (2007). Orientation anisotropies in visual search revealed by noise. *J. Vis.* 7, 11.

Thompson, K. G., and Schall, J. D. (2000). Antecedents and correlates of visual detection and awareness in macaque prefrontal cortex. *Vision Res.* 40, 1523–1538.

Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. *Nature* 381, 520–522.

Viyiani, P., and Swensson, R. (1982). Saccadic eye movements to peripherally discriminated visual targets. *J. Exp. Psychol.* 8, 113–126.

Williams, D. E., and Reingold, E. M. (2001). Preattentive guidance of eye movements during triple conjunction search tasks: The effects of feature discriminability and saccadic amplitude. *Psychon. Bull. Rev.* 8, 476–488.

Williams, D. E., Reingold, E. M., Moscovitch, M., and Behrmann, M. (1997). Patterns of eye movements during parallel and serial visual search tasks. *Can. J. Exp. Psychol.* 51, 151–164.

Williams, L. G. (1966). The effect of target specification on objects fixated during visual search. *Percept. Psychophys.* 1, 315–318.

Wolfe, J. M. (1994). Guided Search 2.0: a revised model of guided search. *Can. J. Exp. Psychol.* 1, 202–238.

Yantis, S., and Jonides, J. (1996). Attentional capture by abrupt onsets: New perceptual objects or visual masking. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 1505–1513.

Yarbus, A. L. (1967). *Eye Movements and Vision.* New York: Plenum Press.

Zelinsky, G. (2008). A theory of eye movements during target acquisition. *Psychol. Rev.* 115, 787–835.

Zelinsky, G. J., and Sheinberg, D. L. (1997). Eye movements during parallel-serial visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 244–262.

Zohary, E., and Hochstein, S. (1989). How serial is serial processing in vision. *Perception* 18, 191–200.