

After-search—visual search by gaze shifts after input image vanishes

Li Zhaoping

Department of Computer Science,
University College London, UK



It has been known that if an image containing many random items suddenly disappears, visual persistence and iconic memory vanish within ~ 0.2 seconds, such that observers are unable to report the identities of the items not graced by attention. We show that effects of the saliency of image items can remain for up to more than a second longer, influencing both eye movements and explicit choice. In our experiment, subjects searched for a moderately salient target among 659 non-targets to report whether the target was in the left or right half of the display, guessing if necessary, while their gaze was tracked. In some trials, the search display was replaced by a mask before their gaze reached the target. We call such trials 'after-search' trials, and the behavior after the mask onset is after-searching. In after-search, subjects made 3.1 ± 2.8 saccades on the mask before reporting. In some after-search trials called 'gaze-arrival' trials, they looked at the location of the vanished target after up to several on-mask saccades; in these trials, the subjects' reports were more accurate. This saliency effect begs the question of what memory, saccade generation, or other neural mechanisms might be responsible.

Keywords: visual search, saliency, saccades, psychophysics, masking, after-search

Citation: Zhaoping, L. (2008). After-search—visual search by gaze shifts after input image vanishes. *Journal of Vision*, 8(14):26, 1–11, <http://journalofvision.org/8/14/26/>, doi:10.1167/8.14.26.

Introduction

If an image containing many items is viewed too briefly, there will be insufficient time for visual attention or gaze to visit all items to store them in the memory before the image disappears. In such a case, it is known that traces of information about the image decay very quickly after the display offset when after-images are prevented. Observers are unable to report the identities of the image items not yet visited by attention if questioned more than about 0.2 seconds after the display offset (Shih & Sperling, 2002; Sperling, 1960). Here we are, of course, referring only to cases in which visual items in the display are unrelated to each other and randomly positioned, such that attending some items in the image does not enable the subjects to infer the identities of items not yet visited by attention.

However, visual inputs contain more than just information about the identities of visual objects, i.e., "what"; they also contain two sorts of "where" information, one is of the information about locations of visual objects, and the other is how salient or conspicuous these locations are to attract visual attention in a stimulus driven or bottom-up manner (Koch & Ullman, 1985; Li, 2002; Zhaoping & May, 2007). In particular, more salient locations attract attention or gaze more strongly when other influences on attentional selection are the same. The "where" information may be obtained before the "what" information (Sagi & Julesz, 1985). The "what" of a visual item is believed to be extracted after attention is paid to its location (Treisman & Gelade, 1980).

One may then ask whether information about the saliencies of locations persists only as briefly as the information about the identities of objects after visual inputs vanish. There are some reasons to think that it might. First, it has been shown that the effect of saliency, manifested by choosing a location to direct saccades or attention, does not last for more than about 0.2 seconds (Nakayama & Mackeben, 1989; van Zoest & Donk, 2006) after stimulus onset, even when the visual display persists. This finding was, however, obtained in studies for which attentional selection was also influenced by top-down or goal-directed control, which has a slower onset and is sustained. Hence, it may have been that the longer-lasting effects of the stimulus-driven, or bottom-up, saliency were simply masked by the effects of the top-down attentional control in these studies. Second, in a study designed to minimize top-down interference, attentional preference for the most, over the second-most, salient item in a display containing other non-salient background items did not last more than about 0.2 seconds (Donk & van Zoest, 2008). However, the relative saliency or the prominence of these two most salient items against the background items could still be longer lasting.

Meanwhile, observations of saccadic planning behavior indicate that the choice of the saccadic destination is typically made afresh during each fixation, or within the most recent inter-saccadic interval (ISI, Findlay, Brown, & Gilchrist, 2001; Gersch, Kowler, & Doshier, 2004; Kowler, Anderson, Doshier, & Blaser, 1995). If the saccadic goal was planned before this, then the resulting ISI is very short, typically below 100 millisecond (ms, McPeck,

Skavenski, & Nakayama, 2000). As each ISI is typically about 300 ms, these observations suggest that saccadic goal is typically determined by the visual information within the most recent 300 or 400 ms, comparable with the known duration of effectiveness of the saliency information. Of course, in assessing saliency persistence and saccadic planning here, we are only considering visual object locations that have not previously been visited by attention, just as in studies of visual persistence and iconic memory (Shih & Sperling, 2002; Sperling, 1960). Once a visual item or location has been attentively viewed, it can be stored in the much longer term memory, such that long after visual inputs vanish, it can give successful recall of the object identity and saccade to the remembered location (Altmann, 2004; Richardson & Spivey, 2000).

In this paper, we present a finding that is unexpected based on these previous results: even after visual input has vanished, aspects of a location within it can continue to affect saccadic behavior and explicit choice in a visual task up to more than 1000 ms, or several saccades, later, even though this location is not salient enough to have attracted overt or covert attention before the image vanished. In fact, because this was unexpected it was discovered only accidentally from the control experimental trials of another study (Zhaoping & Guyader, 2007). It was in a visual search task in which observers searched for a moderately salient target among 659 non-target items in a display spanning $34^\circ \times 46^\circ$ in visual angle, while their gaze was tracked. Observers had to report whether the target was in the left or right half of the display. The target was only weakly salient so that if the search display stayed on, untrained observers would always have to saccade to the target in order to perform the task (Zhaoping & Guyader, 2007). However, in experiments reported here, a mask often replaced the search display before the gaze reached the target. Nevertheless, we show that the information about the target continued to influence the observers' saccadic behavior on the mask and their reports of the target location for up to more than a second longer than the duration of the known visual persistence and saccadic planning. Previous preliminary reports of our findings can be found in Zhaoping (2006) and Zhaoping and Guyader (2006). The implications of this result and its relationship with other studies will be treated in the Discussion section.

Methods

Definitions of terms

For ease of presentation, here we define various terms used throughout the paper. An after-search trial is a visual search trial in which the gaze has not reached the search target before the mask onset, and after-search refers to the

search behavior after the mask onset. Gaze is said to have reached the target when it is within 2.3° in visual angle of the target's center position, and 2.3° is about 1.5 times the average distance between two neighboring search items. A change d in eye position is called a saccade when this change is at least $\Delta \equiv 1.5^\circ$ (the average distance between nearest neighboring items) within 20 ms, the time between two successive samples of the eye tracker. Two such sufficiently large changes d_1 and d_2 in eye positions are defined as two separate saccades if they start at least 100 ms apart from each other, or, if the difference θ in their directions is at least $\phi \equiv 20^\circ$ and $\theta \cdot \min(d_1, d_2)/\Delta > 2\phi$ (where $\min(d_1, d_2)$ is the smaller of d_1 and d_2). All definitions below refer only to events during after-search in the after-search trials which are the concern of this paper. If during an after-search trial, gaze arrives at the location of the vanished target during the after-search, i.e., after the mask onset, this trial is called a gaze arrival trial, or arrival trial for short (see Figure 3 for examples), otherwise, it is a (gaze) non-arrival trial. Even though the target has vanished when the gaze arrives at its location during after-search, for simplicity we still refer to it as gaze reaches or arrives at the target, or gaze arrival. A gaze arrival typically refers to the arrival by gaze to the target for the first time during a trial unless otherwise stated. An after-search saccade is a saccade made during after-search. In an arrival trial, if gaze does not reach the target till after the second or subsequent after-search saccade, we call this a gaze wondering trial, or a wondering trial for short (e.g., Figures 3C–3I), otherwise, it is a (gaze) non-wondering trial (e.g. Figures 3A and 3B). In a small minority of gaze arrival trials in our data, no after-search “saccades” as defined above were made before gaze reached the target, see Figure 4A. This simply means that the gaze moved toward the target so slowly since mask onset that it was not counted as a “saccade” by our definition. The time interval between mask onset and the first gaze arrival in an arrival trial is called the mask-to-arrival latency T_{M-A} , the interval between the gaze arrival and the report by button press is called the arrival-to-report latency T_{A-R} , and the interval between mask onset and the report by button press is called mask-to-report latency T_{M-R} , see Figure 1.

Participants

All observers (subjects) were adults between 18 and 45 years old, had normal or corrected-to-normal vision, and, were naive to the research goal of the experiments.

Stimuli and procedures

Each stimulus, viewed at a distance of 40 centimeters, had $660 = 22 \times 30$ object items, each at a position randomly displaced, up to $\pm 0.24^\circ$ visual angle, horizontally and vertically, from its corresponding position in a

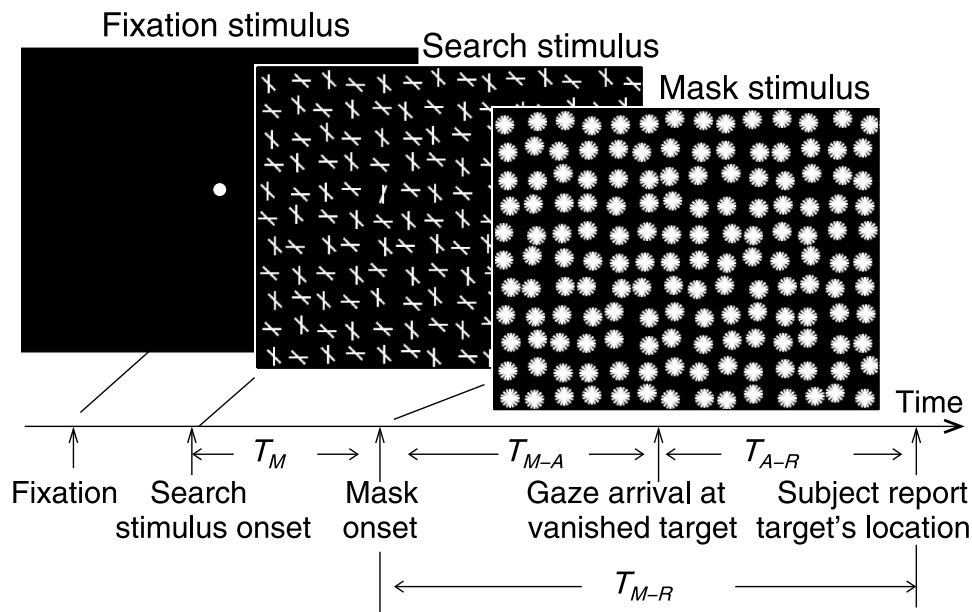


Figure 1. Illustration of the experimental stimuli and events in a gaze arrival after-search trial. The actual stimuli have 660 items, covering $34^\circ \times 46^\circ$ in visual angle. In this example, the search target item is 5th from top and left. All items in the mask stimulus are identical to each other, each replacing one of the object items in the search stimulus, and there is no afterimage of the search stimulus upon masking.

regular grid of 22 rows by 30 columns spanning $34^\circ \times 46^\circ$ in visual angle. Each stimulus bar was $0.12^\circ \times 1.1^\circ$ in visual angle, and 48 cd/m^2 in brightness. The background was black. The target's grid location was randomly one of those closest to the circle of about 15° eccentricity, and beyond 12° of horizontal eccentricity, from the display center. The fixation stimulus was a bright disk of 0.3° diameter at the display center.

Each of the 660 image items had two intersecting bars, one horizontal or vertical, the other oblique. For the 659 non-target items, the oblique bars were 45° from vertical, uniformly tilted left or right for a given display. For the target item, the oblique bar was tilted in the opposite direction from vertical by at least 20° (see Figure 1). The target is salient against the background because it has the uniquely oriented bar in the image (Treisman & Gelade, 1980), but the non-uniformity in the orientations of the background (horizontal and vertical) bars greatly reduces the degree of the target saliency (Duncan & Humphreys, 1989). In trials involving a mask, each item in the search display was replaced by one item in the mask. Mask items were identical, made of eight intersecting bars of the same length and width as the bars in the search items. We verified that upon mask onset, the search stimulus became imperceptible and generated no afterimage.

The observers were told that the target was the item with a uniquely tilted bar. They were told to press a button at their leisure, before or after the mask onset, to report whether the target was in the left or right half of the display, guessing if they had to. Their gaze was tracked by

a 50 Hz infrared video eye tracker from Cambridge Research System (www.crsi Ltd.com) in each trial, from before the stimulus onset till 300 ms after the button press. Tracking calibration was performed before each data session to a precision typically within 0.5° of visual angle. Subjects were instructed to fixate centrally until stimulus onset and to move their eyes freely afterwards to search for the target. Each subject was untrained, and only practiced 2–4 trials immediately before each data-taking session to become familiar with the procedure and the stimuli. The sequence of events in a trial was as follows:

1. With the fixation stimulus, the subject pressed a button to start a trial and eye tracking.
2. After 0.6 second, upon the subject's continuous fixation for 40 ms within 3° of the fixation point, a blank screen replaced the fixation stimulus for 200 ms, followed by the onset (designated as time zero) of the search stimulus.
3. A mask replaced the search stimulus at a time determined as follows.

In a randomly chosen half of the trials, called gaze contingent trials, the mask onset occurred upon or after the first gaze arrival at the target. In the other, non-gaze contingent, trials, a time τ was chosen randomly and uniformly from the time window 200–1700 ms (which was shorter than the average latency needed for gaze to reach the target in an unmasked display). The mask onset occurred upon the first gaze arrival at the opposite side

(laterally from the center) of the target that happened at least 200 ms after stimulus onset, or at time τ , whichever was sooner. Triggering the mask onset by a gaze shift away from the target hemi-field of the display does not evidently give any cues to the subjects regarding the target location, since in such trials only 21% of the first saccades after mask onset were directed to the opposite hemi-field, compared to 29% in the other after-search trials. Most of the gaze-non-contingent trials were after-search trials in which the mask onset occurred before the gaze arrived at the target. The mask, once displayed, remained until after the subject's button press and after the end of eye tracking.

Each data session for each subject had 60 to 130 trials. Each subject participated in no more than three sessions, and performed a total of no more than 260 trials. This was to prevent the subjects from being so over-practiced with the task that they would search for the target without making eye movements. Each subject was asked for his or her comments and observations after each data-taking session. We thereby verified that subjects did not notice any link between mask onset and gaze positions, and regarded the mask onset time as random. Subjects reported that they did not see the target in many trials, and that in these trials they had to guess to make their button press reports.

Data analysis

A trial is defined as being bad and is removed from further analysis if gaze was untracked in more than 10% of the video frames of the eye tracker during the time between stimulus onset and button response, or if the button response occurred within 100 ms after the stimulus onset. Data from a session in which more than 10% of the trials were bad are removed from further analysis. Sufficiently large gaze tracking error can lead to failures to detect gaze arrivals at the target. Hence, if more than 11% of the designated gaze contingent trials in a session were non-arrival trials (that is, the trials ended before the eye tracker's measurement indicated that the gaze has reached the target), poor tracking accuracy is suspected and the data in this session are removed from analysis. Consequently, the final data analysis includes data from 29 sessions by a total of 17 subjects. This gives a total of 1121 after-search trials, of which 148 are after-search gaze arrival trials. Each data session that was included had between 1 to 12 after-search arrival trials (mean 5.1) and each subject that was included contributed 1 to 3 sessions and between 1 to 17 after-search gaze arrival trials. The statistics are based on pooling these after-search trials from all included subjects and sessions.

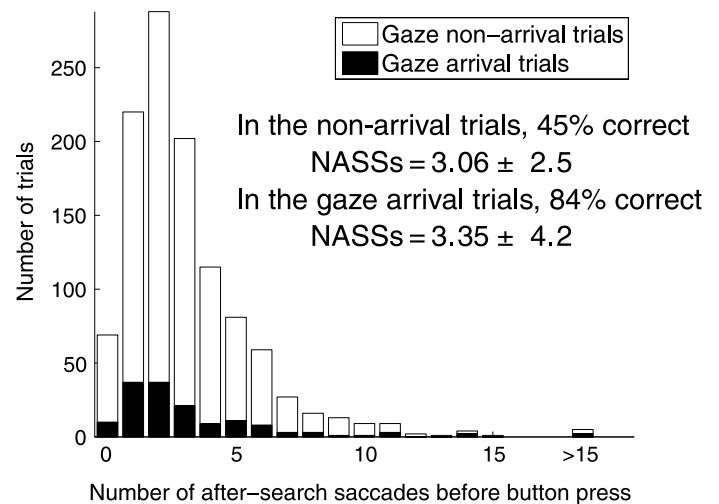
All significance tests were based on two tailed t -test. The errors attributed to the means of data are standard errors of the mean, while the errors for individual data points are standard deviations.

Results

All the results presented concern only the after-search trials except when explicitly pointed out otherwise. The findings on the non-after-search trials (the gaze contingent trials) have been reported elsewhere (Zhaoping & Guyader, 2007).

After mask onset in the after-search trials, subjects waited for a latency of $T_{M-R} = 1.1 \pm 0.8$ seconds before they pressed the button to report the target location, even though they could choose to report immediately after the

A: Number of after-search saccades (NASSs) before button press



B: Mask-to-report latency T_{M-R}

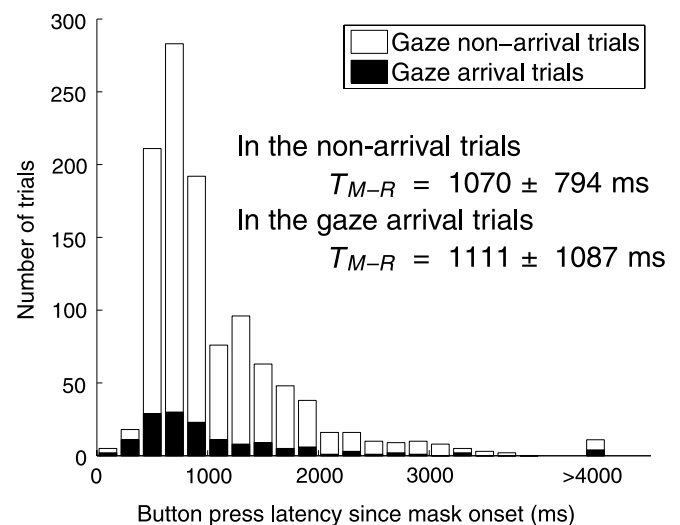


Figure 2. Distributions of the numbers of after-search saccades (NASSs) before button press (A) and of the latencies T_{M-R} of the button presses during after-search (B) among 1121 after-search trials. There were 148 after-search arrival trials (shown in the filled part of the bars) in which gaze arrived at the target during after-search. Button press report was correct in 84% of the gaze arrival trials and 45% of the gaze non-arrival trials.

mask onset, see Figure 2. During this button press latency T_{M-R} , they made 3.1 ± 2.8 after-search saccades on the mask. In most of the after-search trials, gaze never

reached the target before button press. However, 148 of the 1121 after-search trials are gaze arrival trials in which the after-search eye movements brought the gaze to the

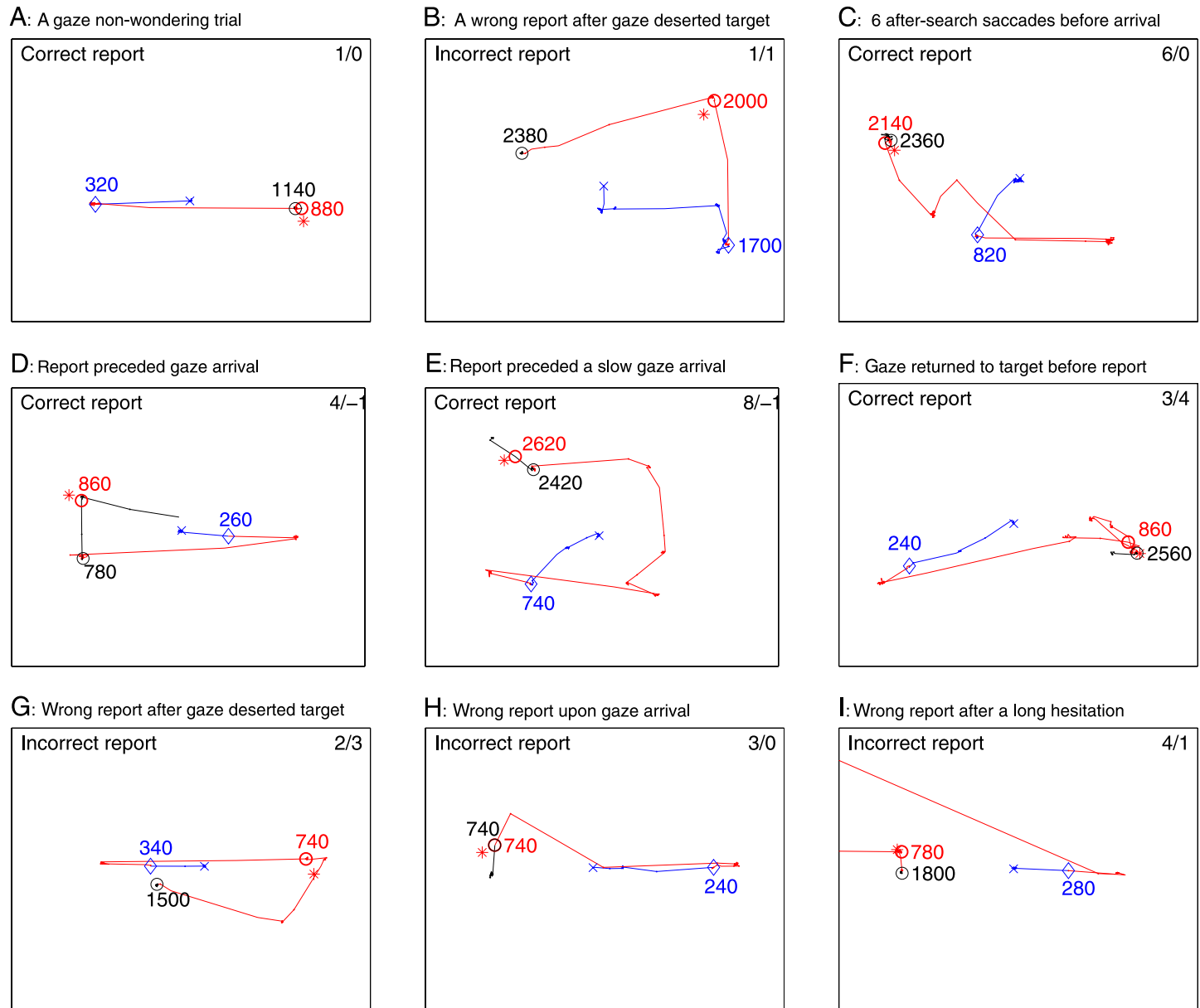


Figure 3. Nine examples of gaze arrival after-search trials to illustrate some possible behaviors in these trials. Each plot frames the visual space containing the 660 visual items, the target location is marked by a red “*”, and a blue cross marks the eye position at search stimulus onset. Blue traces mark eye positions before mask onset, red ones mark eye positions between mask onset and button press report, and black ones for eye positions during the 300 ms after the report. Blue diamond, red circle, or black circle mark the eye position at the mask onset, when gaze was judged as arrived at the target for the first time, or at button press respectively. Next to these marks are numbers in corresponding colors to indicate the latency (in ms) of these corresponding events since stimulus onset. The accuracy of the button press (correct or incorrect) is marked at the top left corner of each plot. The numbers n_1 and n_2 on the top right corner, in n_1/n_2 , indicate the numbers of after-search saccades between the mask onset and the gaze arrival, and between the gaze arrival and the button press respectively. (A) and (B) are gaze non-wondering trials in which the gaze reached the target by no more than one after-search saccade, and the rest are gaze wondering trials in which more than one after-search saccades were needed for gaze to reach target. In (A), (C), and (H), the button was pressed without hesitation and without any additional after-search saccades following the arrival of the gaze at the target; in (D) and (E), the button was pressed one saccade before the gaze arrival at the target. In the rest of the examples, the button press occurred only after hesitation, and after at least one additional after-search saccade following the gaze arrival.

target (see Figure 3 for examples). If any gaze were directed randomly, with equal chance, to each unit area of the $46^\circ \times 34^\circ$ on the stimulus display, then 3.1 saccades in a trial would give a 3.29% chance for the gaze to reach the target in a trial (i.e., to fall within 2.3° from the center of the target), there would be only on average 37 arrival trials out of the 1121 after-search trials, and the chance to have at least 148 arrival trials out of the 1121 after-search trials would be $p < 10^{-44}$ (calculated from a binomial distribution of the number of arrival trials based on each trial's chance of target arrival). Furthermore, the task performance was correct in 84% of the gaze arrival trials but was near chance in the gaze non-arrival trials. Given each subject s (for $s = 1$ to 17 subjects) had N_s arrival trials, if one randomly draws N_s trials from all the after-search trials by the corresponding subject and pools these trials across subjects, then the chance to have at least 84% correct responses would be $p < 0.0001$ (calculated from a Monte Carlo simulation of such a random sampling process). Hence, the gaze arrival was not by chance and had behavioral consequences. Meanwhile, if we compare the gaze arrival and non-arrival trials, there is no significant difference in the numbers of after-search saccades ($p = 0.24$) or the button press latency T_{M-R} ($p = 0.57$) following mask onset.

Figure 3 shows some examples of the after-search arrival trials to illustrate some possible behaviors in such trials. Note that the button press often followed very soon after, or sometimes occurred briefly before, the gaze arrived at the target for the first time, for example in Figures 3A, 3C, 3D, 3E, and 3H. By comparison, in the trials of Figures 3B, 3F, 3G, and 3I, the button press did not occur until after at least one additional saccade after the gaze reached the target. In Figures 3D and 3E, the button press occurred slightly before gaze arrival, as if the hand response anticipated the gaze arrival. In Figures 3C and 3E, it took especially many, 6 and 8 respectively, after-search saccades for the gaze to reach the target, but the button press occurred immediately after or before the long delayed arrival of the gaze. In Figure 3H, the button press and gaze arrival occurred simultaneously, even though the arrival took 3 after-search saccades; nevertheless a wrong button was pressed, although such cases are rare, see Figures 4A and 4C. Figure 3I shows an example in which it took four after-search saccades to reach the target, including a period in which gaze briefly strayed outside the display screen, and then there followed an unusually long (1020 ms) hesitation during which the gaze stayed close by the target, before the wrong button was finally pressed. In the after-search in Figure 3F, the gaze arrived, departed, and then returned to the target before a correct button press. In Figures 3B and 3G, a wrong button response was preceded by the gaze deserting the target after its arrival in after-search.

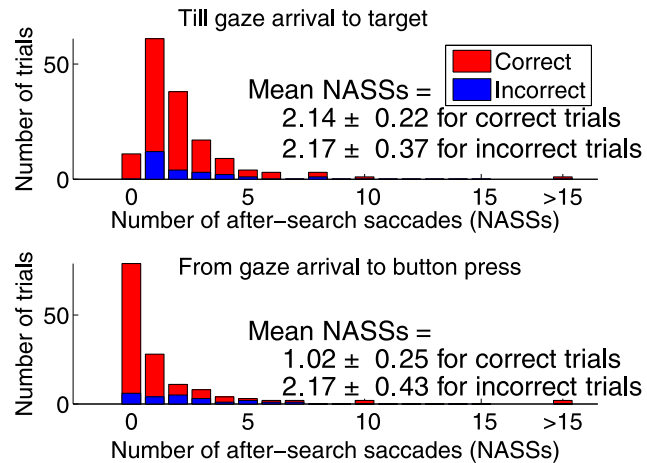
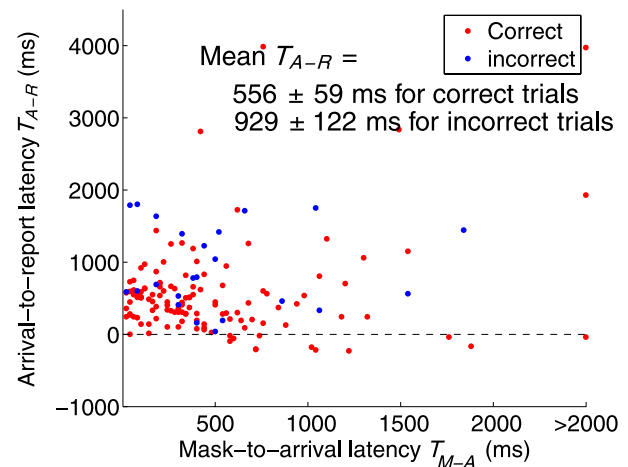
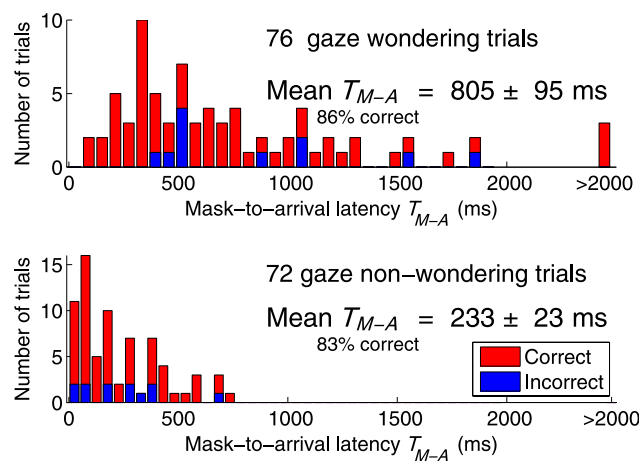
As a saccade is typically planned within the most recent inter-saccadic interval (ISI), it is possible that in some of the after-search trials, the first after-search saccade was

planned just before the mask onset. If subjects were more aware of the target in the gaze arrival trials which required fewer after-search saccades to reach target, then one should see an inverse correlation between task performance and the number of after-search saccades needed to reach the target. However, Figure 4A shows that, in a gaze arrival after-search trial, it requires an average of 2.15 after-search saccades to reach the target, regardless of the task report accuracy ($p = 0.96$). Figure 4B shows that the task performance in the gaze wondering trials, which required more than one after-search saccades for gaze arrival to target, is no worse than that in the non-wondering trials, even though the gaze arrival latency T_{M-A} (on average 805 ± 96 ms) of the former is significantly ($p < 0.001$) longer. This suggests that subjects were no more aware of the target in the arrival trials requiring fewer after-search saccades to reach target than they were in the other after-search arrival trials.

Meanwhile, the correctly and incorrectly reported after-search arrival trials differ in behavior after the gaze arrived at the target. In particular, the incorrect trials required marginally more (on average 2.17 ± 0.43) after-search saccades between the gaze arrival and button press report ($p = 0.059$, Figure 4A), and significantly ($p = 0.012$, Figure 4C) longer report latency T_{A-R} (on average 929 ± 122 ms) since gaze arrival. Hence, the incorrect arrival trials tend to be more hesitantly reported after the gaze arrival during after-search.

While most saccades are planned afresh at each fixation, it is known that some saccades can be planned before the most recent fixation (McPeck et al., 2000). Thus, the second after-search saccade after mask onset could possibly be planned before the mask onset. However, a saccade planned before the most recent fixation is typically preceded by a very short inter-saccadic interval (ISI) typically below 100 ms (McPeck et al., 2000). In contrast, most of the ISIs between our after-search saccades are longer than 200 ms (see Figure 4D), among the gaze arrival and non-arrival trials. Hence, the second and subsequent after-search saccades, some of them leading to the gaze arrival to target with a consequently superior task performance, are most likely planned after the mask onset, i.e., not due to a covert attentional shift to target before the mask onset. It may still be possible that a saccade planned to the target, even to arrive after a transitory fixation somewhere else, was interrupted by the onset and the continued presence of the mask, thereby losing the characteristic short ISI if there is any. To investigate this, ISIs across the mask onset, each between a saccade before and another after the mask onset, were measured and found to have a mean ISI = (367 ± 10) ms. Meanwhile, ISIs between saccades before target arrival in the non-after-search trials (i.e., the gaze-contingent trials in which the mask onset occurred after gaze arrival) were found to have a mean ISI = (369 ± 7) ms (this is calculated by pooling the randomly sampled gaze-contingent trials across subjects, each contributing a share of the trials

A: Number of after-search saccades (NASSs)

C: Arrival-to-report latencies T_{A-R} B: Mask-to-arrival latencies T_{M-A} 

D: Inter-Saccadic Intervals (ISIs)

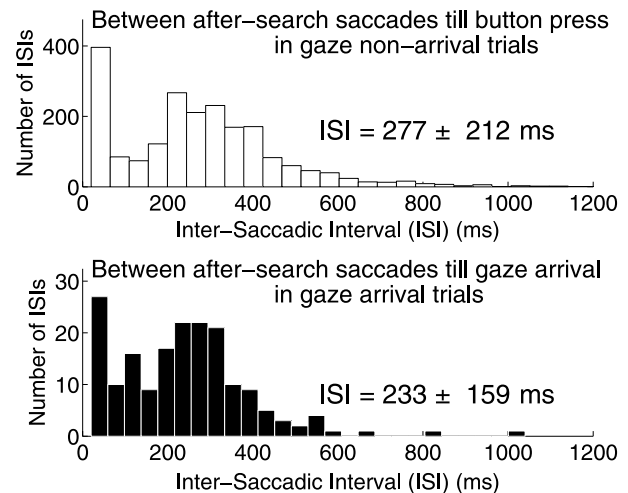


Figure 4. Statistics of behavior during after-search. In (A)–(C), (trials with) correct button presses are indicated in red; incorrect presses in blue. (A) Number of after-search saccades (NASSs) in gaze arrival trials separated according to those before the arrival of the gaze to the target (upper) and after gaze arrival, but before button response (lower). Note that NASSs before the gaze arrival are no different ($p = 0.96$) in correct and incorrect trials, but the incorrect trials have marginally ($p = 0.059$) more after-search saccades between gaze arrival and button press. (B) The latency T_{M-A} of gaze arrival following mask onset in the gaze wandering trials (upper) is significantly ($p < 0.001$) longer than that in the non-wondering trials (lower), although the task performance is no worse in the gaze wandering trials. (C) Button press report latency T_{A-R} since the gaze arrival versus the gaze arrival latency T_{M-A} since mask onset in the arrival trials, T_{A-R} is significantly ($p = 0.012$) longer in the incorrect than the correct trials. (D) Inter-saccadic-intervals (ISI) between after-search saccades for gaze arrival and non-arrival trials.

matching his/her share in the pooled after-search trials). These two groups of ISIs are indifferent from each other ($p = 0.90$), suggesting that the saccades planned just before the mask onset in the after-search trials were not significantly disturbed by the onset of the mask.

Discussion

Human observers have a tendency to look back at the locations of the previously viewed visual objects when

they recall or listen to information associated with those objects, even after these objects have vanished (Altmann, 2004; Richardson & Spivey, 2000). These findings indicate that the locations of the inputs that have been processed by attention are present in memory, and ocular-motor coordinates maybe used by the brain as pointers to memories of objects or events. Physiological observations suggest that this memory may reside in the intra-parietal areas of the brain, which are associated with attentional or gaze control and sensory-motor transformations (Bisley & Goldberg, 2003; Mazzoni, Bracewell, Barash, & Anderson 1996; Schall & Thompson, 1999).

It is known (Hoffman, 1998; Kowler et al., 1995) that gaze is mandatorily directed to the focus of attention in tasks such as ours in which it is not constrained. Further, gaze follows spatial shifts of this focus closely in time. In our after-search trials, the gaze had not reached the target before the search display disappeared. Hence, unless a saccade to the target was being planned just before the mask onset, the target had not been visited by (covert) attention before it disappeared. In the gaze non-wondering after-search arrival trials in which gaze reached the target following no more than one after-search saccade, it is likely that these saccades were planned before mask onset or that attention had covertly shifted to the target. However, from the analysis above of the ISIs in our data, the second or subsequent after-search saccades in the after-search trials were unlikely to have been planned before the offset of the search display. Hence, in the gaze-wondering after-search arrival trials, in which gaze reached target following more than one after-search saccade, it is unlikely that attention had covertly visited the target before it disappeared, although the data analysis can not conclusively rule out this possibility. Therefore, I suggest that the after-search saccades to the target in these trials are of a different kind from saccades to previously viewed and remembered locations in the previous studies (Altmann, 2004; Richardson & Spivey, 2000).

If the second and subsequent after-search saccades to the target are planned after the search display vanishes, they would have to be planned from some sort of memory of the target. If attention had not visited the target, the memory responsible for these after-search saccades must therefore be pre-attentive. It is likely that such a memory is that of a pre-attentive or bottom-up saliency map of the visual input. Even though saliency does not have the “what” information about object identities, it carries the information about the conspicuousness of the visual locations, and is therefore sufficient for our task which only requires reporting the location and not the identity of a moderately salient input item. Though iconic memory and visual persistence about the identity of visual objects last for no more than ~0.2 seconds (Shih & Sperling, 2002; Sperling, 1960), we could account for our observations by hypothesizing that memory of the saliencies of visual locations lasts longer. As this memory survives gaze shifts during after-search, it should thus be encoded in a world-centered rather than retinally centered coordinate system. Experimental observations (Allman, Miezin, & McGuinness, 1985; Jingling & Zhaoping 2008; Jones, Grieve, Wang, & Sillito, 2001; Knierim & van Essen, 1992; Koene & Zhaoping, 2007; Lamme, 1995; Li & Li, 1994; Nothdurft, Gallant, & Van Essen, 1999; Sillito, Grieve, Jones, Cudeiro, & Davis, 1995; Wachtler, Sejnowski, & Albright, 2003; Zhaoping, 2008; Zhaoping & May, 2007; Zhaoping & Snowden 2006) as well as theoretical and modeling analysis (Li, 1999a, 1999b, 2000, 2002) have suggested that the retinotopic primary visual cortex (V1) creates this bottom-up saliency

map, particularly for saliency due to contrasts in the low level orientation features in visual inputs, as is the case in our stimuli. The information in this map, after the associated visual inputs had vanished in our trials, with activities in V1 presumably being replaced by those evoked by the mask, might be stored in the superior colliculus or lateral-intra-parietal cortex (Carello & Krauzlis, 2004; Krauzlis & Dill, 2002; Schiller, 1998; Shadlen & Newsome, 2001; Tehovnik, Slocum, & Schiller, 2003) both of which receive inputs from V1, direct saccades, and exhibit neural activities that are correlated with cognitive decisions. The following observations may provide clues to the underlying mechanisms for this memory and its function. The ISIs (<300 ms on average) during after-search (Figure 4D) is much shorter than the ISIs (>300 ms on average) before the target arrival (and thus before the mask onset) in the non-after-search trials ($p < 0.0001$, here the comparison is not made with the pre-mask ISIs in the after-search trials since the earlier mask onset biases the sampled ISIs to shorter ones). Furthermore, the reaction time RT_{eye} for gaze arrival to target (since stimulus onset) in the after-search arrival trials is on average (1137 ± 62) ms, significantly ($p < 10^{-6}$) shorter than the average $RT_{eye} = (3026 \pm 373)$ ms of the non-after-search trials, pooled across subjects by randomly sampling as many non-after-search trials by each subject as his/her after-search arrival trials. However, if for each subject, one selectively avoid sampling from the non-after-search trials in which RT_{eye} is longer than the subject’s median RT_{eye} (among non-after-search trials), then RT_{eye} of the after-search arrival trials are not significantly different ($p = 0.64$) from that in these faster non-after-search trials.

Previous studies have suggested that the effect of bottom-up attraction to attention, i.e., of the bottom-up saliency map, is transient, lasting no more than ~0.2 seconds (Nakayama & Mackeben, 1989) and affecting no more than the first saccade after the stimulus onset (van Zoest & Donk, 2006). These findings were obtained when the effect of the bottom-up saliency was compared with that arising from top-down factors, such as prior knowledge of the likely location of the target, or the resemblance between visual inputs and an internal template of the target. The top-down effects, which are long lasting, may add to, or be in conflict with, the bottom-up effect, thereby masking or removing the manifestation of the bottom-up effect after the initial transients. In our experiment, there was no top-down cue about the likely locations of the target beyond a probabilistic chance level, and, once the search stimulus vanished, any resemblance between the search display and any internal templates should vanish as quickly as the visual persistence for the “what” information. The long lasting persistence of the bottom-up saliency may thereby be uncovered in our experiment by masking the “what” information, and thereby any related top-down effect. Our finding also does not contradict the finding by Donk and

van Zoest (2008) that the effect of the relative saliency between the two most salient items in the display is short lived even in situation with minimized top-down influences. This is because, even though the relative saliency between the two most salient items maybe short lived, the observers were still able to saccade to either of them among the non-salient background items after a longer delay, suggesting that the information about the presence of salient location(s) or visual objects (whatever their identity) against a background could persist longer.

As the first finding of a most unexpected phenomenon, our study leaves many additional questions about the long-lasting persistence of bottom-up saliency unanswered. For instance, our subjects reported that they did not see the target in many trials and had to guess. Thus, they may not be aware that some of their gaze shifts and reports are more accurate than they thought. This is consistent with recent findings that attentional shifts driven by bottom-up saliency can be dissociated from awareness (Zhaoping, 2008), and that visual stimuli invisible to awareness can affect visual task performance (Jiang, Costello, Fang, Huang, & He, 2006) and activate brain areas processing “where” information for visually guided actions (Fang & He, 2005). However, it would be good to probe explicitly whether observers are aware of the target in the after-search gaze arrival trials. It is likely that observers did not (covertly) see the target before the mask onset in the gaze-wondering arrival trials, otherwise it would be puzzling why they did not saccade to the target immediately after the mask onset, and in some trials wasted several wrong saccades before arriving at the target. Conversely, if, in fact, the observers did “see” the target in the gaze wondering after-search trials before they could saccade to it prior to its disappearance, it would imply that human saccades could in fact be planned further in the future, and by many more saccades in advance than previously believed. One may also ask how and which neural structures store the visual information, surviving the dynamic gaze shifts, and more generally how the interactions between the sensory and motor structures serve the perceptual and cognitive functions that we observed.

Conclusion

It is reported that the effect of bottom-up visual saliency about visual locations can persist more than one second longer than the ~0.2 second persistence of visual identity information. This persistence of saliency is manifested in saccadic behavior and cognitive choice of locations after a visual stimulus has disappeared, and may be revealed when the competing top-down influences on behavior are absent or are less dominant. These findings should motivate future investigations in, e.g., memory, the interactions between sensory and motor functions, and the underlying neural mechanisms.

Acknowledgments

This research was supported by a grant from the Gatsby Charitable Foundation and a Cognitive Science Foresight grant BBSRC #GR/E002536/01. I thank Nathalie Guyader for her help in collecting the data, and two anonymous reviewers for comments.

Commercial relationships: none.

Corresponding author: Li Zhaoping.

Email: z.li@ucl.ac.uk.

Address: Department of Computer Science, University College London, Gower St., London WC1E 6BT, UK.

References

- Allman, J., Miezin, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local–global comparisons in visual neurons. *Annual Review of Neuroscience*, 8, 407–430. [[PubMed](#)]
- Altmann, G. T. (2004). Language-mediated eye movements in the absence of a visual world: The ‘blank screen paradigm’. *Cognition*, 93, B79–B87. [[PubMed](#)]
- Bisley, J. W., & Goldberg, M. E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science*, 299, 81–86. [[PubMed](#)]
- Carello, C. D., & Krauzlis, R. J. (2004). Manipulating intent: Evidence for a causal role of the superior colliculus in target selection. *Neuron*, 43, 575–583. [[PubMed](#)] [[Article](#)]
- Donk, M., & van Zoest, W. (2008). Effects of salience are short-lived. *Psychological Science*, 19, 733–739. [[PubMed](#)]
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458. [[PubMed](#)]
- Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, 8, 1380–1385. [[PubMed](#)]
- Findlay, J. M., Brown, V., & Gilchrist, I. D. (2001). Saccade target selection in visual search: The effect of information from the previous fixation. *Vision Research*, 41, 87–95. [[PubMed](#)]
- Gersch, T. M., Kowler, E., & Doshier, B. (2004). Dynamic allocation of visual attention during the execution of sequences of saccades. *Vision Research*, 44, 1469–1483. [[PubMed](#)]
- Hoffman, J. E. (1998). Visual attention and eye movements. In H. Pashler (Ed.), *Attention* (pp. 119–154). San Diego: Psychology Press.

- Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 17048–17052. [PubMed] [Article]
- Jingling, L., & Zhaoping, L. (2008). Change detection is easier at texture border bars when they are parallel to the border: Evidence for V1 mechanisms of bottom-up salience. *Perception*, 37, 197–206. [PubMed]
- Jones, H. E., Grieve, K. L., Wang, W., & Sillito, A. M. (2001). Surround suppression in primate V1. *Journal of Neurophysiology*, 86, 2011–2028. [PubMed] [Article]
- Knierim, J. J., & van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, 67, 961–980. [PubMed]
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, 4, 219–227. [PubMed]
- Koene, A. R., & Zhaoping, L. (2007). Feature-specific interactions in salience from combined feature contrasts: Evidence for a bottom-up saliency map in V1. *Journal of Vision*, 7(7):6, 1–14, <http://journalofvision.org/7/7/6/>, doi:10.1167/7.7.6. [PubMed] [Article]
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35, 1897–1916. [PubMed]
- Krauzlis, R., & Dill, N. (2002). Neural correlates of target choice for pursuit and saccades in the primate superior colliculus. *Neuron*, 35, 355–363. [PubMed] [Article]
- Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, 15, 1605–1615. [PubMed] [Article]
- Li, C. Y., & Li, W. (1994). Extensive integration field beyond the classical receptive field of cat's striate cortical neurons—Classification and tuning properties. *Vision Research*, 34, 2337–2355. [PubMed]
- Li, Z. (1999a). Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 10530–10535. [PubMed] [Article]
- Li, Z. (1999b). Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network*, 10, 187–212. [PubMed]
- Li, Z. (2000). Pre-attentive segmentation in the primary visual cortex. *Spatial Vision*, 13, 25–50. [PubMed]
- Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, 6, 9–16. [PubMed]
- Mazzoni, P., Bracewell, R. M., Barash, S., & Andersen, R. A. (1996). Motor intention activity in the macaque's lateral intraparietal area. I. Dissociation of motor plan from sensory memory. *Journal of Neurophysiology*, 76, 1439–1456. [PubMed]
- McPeck, R. M., Skavenski, A. A., & Nakayama, K. (2000). Concurrent processing of saccades in visual search. *Vision Research*, 40, 2499–2516. [PubMed]
- Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, 29, 1631–1647. [PubMed]
- Nothdurft, H. C., Gallant, J. L., & Van Essen, D. C. (1999). Response modulation by texture surround in primate area V1: Correlates of “popout” under anesthesia. *Visual Neuroscience*, 16, 15–34. [PubMed]
- Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: Looking at things that aren't there anymore. *Cognition*, 76, 269–295. [PubMed]
- Sagi, D., & Julesz, B. (1985). “Where” and “what” in vision. *Science*, 228, 1217–1219. [PubMed]
- Schall, J. D., & Thompson, K. G. (1999). Neural selection and control of visually guided eye movements. *Annual Review of Neuroscience*, 22, 241–259. [PubMed]
- Schiller, P. (1998). The neural control of visually guided eye movements. In J. E. Richards (Ed.), *Cognitive neuroscience of attention* (pp. 3–50). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, 86, 1916–1936. [PubMed] [Article]
- Shih, S. I., & Sperling, G. (2002). Measuring and modeling the trajectory of visual spatial attention. *Psychological Review*, 109, 260–305. [PubMed]
- Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature*, 378, 492–496. [PubMed]
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74, 1–29.
- Tehovnik, E. J., Slocum, W. M., & Schiller, P. H. (2003). Saccadic eye movements evoked by microstimulation of striate cortex. *European Journal of Neuroscience*, 17, 870–878. [PubMed]
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136. [PubMed]
- van Zoest, W., & Donk, M. (2006). Saccadic target selection as a function of time. *Spatial Vision*, 19, 61–76. [PubMed]

- Wachtler, T., Sejnowski, T. J., & Albright, T. D. (2003). Representation of color stimuli in awake macaque primary visual cortex. *Neuron*, 37, 681–691. [[PubMed](#)] [[Article](#)]
- Zhaoping, L. (2006). Visual search with and without target object recognition. *Perception* 35 ECVF Abstract Supplement. [[Article](#)]
- Zhaoping, L. (2008). Attention capture by eye of origin singletons even without awareness—A hallmark of a bottom-up saliency map in the primary visual cortex. *Journal of Vision*, 8(5):1, 1–18, <http://journalofvision.org/8/5/1/>, doi:10.1167/8.5.1. [[PubMed](#)] [[Article](#)]
- Zhaoping, L., & Guyader, N. (2006). Blind search—Successful saccades to the unknown target location up to 1000 ms after removal of visual search stimulus [[Abstract](#)]. *Journal of Vision*, 6(6):484, 484a, <http://journalofvision.org/6/6/484/>, doi:10.1167/6.6.484.
- Zhaoping, L., & Guyader, N. (2007). Interference with bottom-up feature detection by higher-level object recognition. *Current Biology*, 17, 26–31. [[PubMed](#)] [[Article](#)]
- Zhaoping, L., & May, K. A. (2007). Psychophysical tests of the hypothesis of a bottom-up saliency map in primary visual cortex. *PLoS Computational Biology*, 3, e62. [[PubMed](#)] [[Article](#)]
- Zhaoping, L., & Snowden, R. J. (2006). A theory of a saliency map in primary visual cortex (V1) tested by psychophysics of color-orientation interference in texture segmentation. *Visual Cognition*, 14, 911–933.