

# Gaze capture by eye-of-origin singletons: Interdependence with awareness

Li Zhaoping

Department of Computer Science,  
University College London, UK



Where we look in visual tasks is determined by both bottom-up and top-down factors. One theory (Li, 1999a, 2002) suggests that visual area V1 creates a bottom-up saliency map, guiding gaze through extensive projections to the superior colliculus. V1 is the only visual cortical area that represents the eye of origin of an input and is also least associated with awareness; I therefore predicted that an ocular singleton (i.e., an item only shown to one eye among other items shown to the other eye) that is perceptually indistinct might nevertheless attract gaze. In visual searches for an orientation singleton target bar among uniformly oriented background bars, an ocular singleton non-target bar, at the same eccentricity as the target from the center of the search display, often captured the first search saccade. The chance of this capture was above 50% (e.g., 75%) when the eccentricity of the singletons was large and luminance did not vary between the bars, and it was below 50% when the eccentricity was smaller and luminance varied. After each search trial, observers reported whether an ocular singleton non-target (which was actually presented in half of the trials) had been shown. When different bars had similar luminance, misses numbered less than 50% and were independent of whether the gaze was captured by the ocular singleton. However, when luminance varied sufficiently between the bars, 50% were missed overall, albeit significantly less for those that captured gaze. The experiments in this work followed the guidelines in the Declaration of Helsinki.

Keywords: gaze capture, eye of origin, singleton

Citation: Zhaoping, L. (2012). Gaze capture by eye-of-origin singletons: Interdependence with awareness. *Journal of Vision*, 12(2):17, 1–22, <http://www.journalofvision.org/content/12/2/17>, doi:10.1167/12.2.17.

## Introduction

Spatial attention in vision is the process by which a small region in the visual field is selected for detailed processing. Attention is necessary because the brain has limited resources for processing visual stimuli. Indeed, detection and discrimination performance are improved at selected compared with non-selected locations, a phenomenon known as the cueing effect of visual attentional selection (Posner & Petersen, 1990). Attentional selection is an act of looking; although it normally involves overtly shifting the gaze, covert shifts of the gaze of the mind's eye are also possible (Hoffman, 1998) and can also be considered as actions. Post-selectional processing leads to seeing or perceiving. Hence, simply put, vision looks and then sees.

The distinction between these two processes, which is also enshrined in the fact that separate brain areas are associated with the motor action of looking and the perceptual process of seeing, suggests the highly counter-intuitive possibility that seeing a visual feature might not necessarily follow from looking at it as the result of its gaze attraction. A loose link between gaze action and perceptual awareness is also suggested by the observation that human observers can follow a target's movements more accurately by their gaze or hand than by their perception (Belopolsky, Kramer, & Theeuwes, 2008; Frith, Blakemore, & Wolpert, 2000; Leach & Carpenter, 2001; Masson, Busettini, &

Miles, 1997; Tavassoli & Ringach, 2010). However, subjects in those studies did have some awareness of the presence of the targets or distractors of their motor actions, and no previous study has reported that subjects were blind to a visual feature when, or after, gaze had been attracted to it. Nevertheless, very recent observations (Zhaoping, 2008a) have tentatively suggested a class of stimuli that lead to looking without seeing. In this paper, I use these stimuli to illuminate aspects of both processes.

In many cases, looking is goal-dependent or intentional (Corbetta & Shulman, 2002; Desimone & Duncan, 1995; Kastner & Ungerleider, 2000; Treisman & Gelade, 1980), for instance, when we focus our gaze on a book while reading. However, looking can also be goal-independent or bottom-up, such as when a sudden movement in the visual periphery distracts attention away from reading. This is often more potent and faster than goal-dependent looking (Jonides, 1981; Nakayama & Mackeben, 1989; Theeuwes, 1992; Theeuwes, Kramer, Hahn, & Irwin, 1998; Yantis, 1998) and is my focus. Here, I refer to the strength with which a location attracts bottom-up selection as its saliency. My question then becomes whether there are locations whose visual features make them highly salient and so readily attract looking but for which those very same visual features are imperceptible.

Locations are typically salient when they differ from their spatiotemporal neighbors in a feature such as color, orientation, brightness, or motion direction. Human subjects are perceptually sensitive to all these features, which

is why it is intuitive to expect a link between saliency and awareness. By contrast, the feature employed by Zhaoping (2008a) involved the eye of origin, a quantity that turned out to lead to high saliency but low awareness. Specifically, it was found that observers could find a visual search target more quickly in the dichoptic congruent (DC) case when the target was presented to one eye and the non-targets to the other eye (see Figure 1B) compared to the monocular (M) case when all items were shown to the same eye (see Figure 1A), even though observers typically did not spontaneously notice any perceptual difference between these two cases (Zhaoping, 2008a). This occurred both when the target was a non-salient “T” among non-target “L”s or when the target was a highly salient orientation singleton (a bar oriented  $50^\circ$  away from uniformly oriented background bars that were otherwise identical to it; see Figure 1). Furthermore, in the dichoptic incongruent (DI) condition, when the eye of origin or ocular singleton was one of the background non-targets (see Figure 1C), it took almost 200 ms longer to find the orientation singleton target. Since 200 ms is comparable to an inter-saccadic interval in a speeded visual search, the observation suggests that the ocular singleton was more salient than the orientation singleton, thereby distracting gaze away from the target. Previously, the only task-irrelevant stimuli that had been known to be able to capture gaze away from a visual search target had themselves been highly perceptually distinctive, for instance, a color singleton non-target impeding search for a uniquely shaped item (Theeuwes, 1992).

The relationship between saliency-evoked looking and subsequent awareness could cast particular light on the neural substrates of saliency. Looking and seeing are more likely to be separate if the substrates of saliency are far from the brain areas known to be involved in perception and awareness of sensory signals. In this context, it is worth noting that the neural activities in the primary visual cortex (V1) are least correlated with perception among all visual cortical areas; it is the more frontal and anterior cortical areas that are associated with perception and awareness (Crick & Koch, 1995). Indeed, Zhaoping’s (2008a) results were predicted by the recent theory (Li, 1999a, 2002) that the saliency of a location is defined by the highest V1 neural response to that location. Uniquely among visual cortical areas, V1 remains sensitive to the eye of origin of a visual input. Thus, it could signal the high saliency associated with the dichoptic congruent and incongruent stimuli, even though the subjects might be wholly unaware of the underlying feature. It could mediate looking via its direct inputs to the superior colliculus, which drives eye movements (Schiller, 1998). Suggestions that saliency is computed in higher brain areas (Itti & Koch, 2001), such as the parietal (Bisley & Goldberg, 2010; Gottlieb, Kusunoki, & Goldberg, 1998) or frontal areas (Serences & Yantis, 2007; Thompson & Bichot, 2005), have more difficulty accounting for Zhaoping’s (2008a) result.

There have also been observations that other subliminal or invisible visual stimuli, such as weak and brief stimuli, attract attention, as manifested in cueing effects to

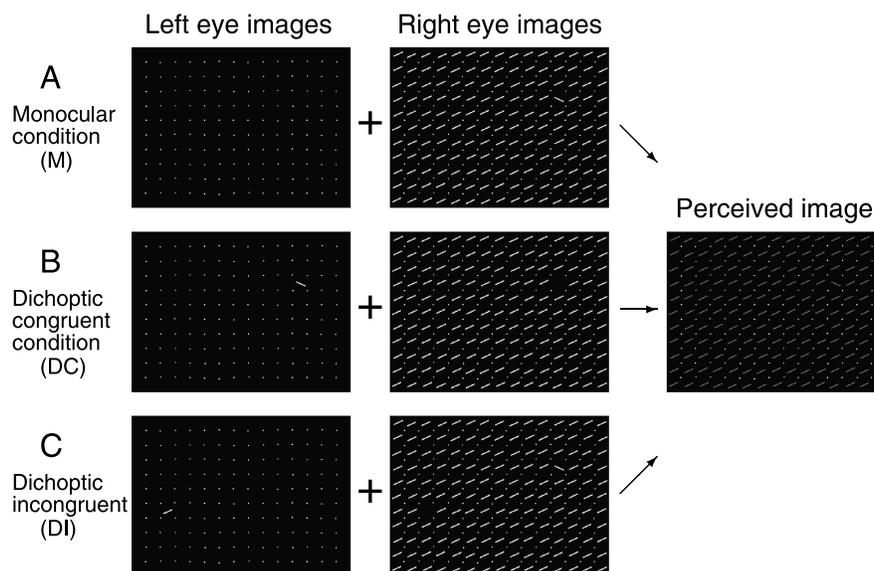


Figure 1. Reduced-size versions of sample visual search stimuli used in Experiment 1a. This studied how a task-irrelevant ocular singleton bar attracts attention in a visual search for an orientation singleton target. The three different dichoptic presentations, monocular (M), dichoptic congruent (DC), and dichoptic incongruent (DI), are the same when left and right eye images are superposed (resembling the perceived image). An ocular singleton bar was absent in the M condition and was the target in the DC condition. In the DI condition, it was always a background bar, with the same eccentricity as the target, but in the opposite lateral half of the perceived image from the target. Observers were asked to report (by pressing a button) as soon as possible whether the target was in the left or right half of the perceived image. The dichoptic condition and the eye of origin of the target were random in each trial.

improve performance or shorten reaction times (RTs) for tasks whose targets appear at their locations (e.g., Jiang, Costello, Fang, Huang, & He, 2006; McCormick, 1997). Although these stimuli attract attention much more weakly than the ocular singleton, these observations support the idea that selection may start from V1, which responds to invisible stimuli.

Zhaoping (2008a) only explored a restricted set of questions associated with eye-of-origin singletons. Here, I consider a broader array. In particular, it is known that a very salient non-target can capture gaze away from a visual search target when the search display is presented for so long that saccades are evoked (Theeuwes et al., 1998). Thus, if the task-irrelevant ocular singleton is indeed more salient than the orientation singleton target, it should be more likely to attract the first saccade in a speeded search (van Zoest & Donk, 2006). This would at least be the case when bottom-up saliency dominates top-down task-dependent factors in guiding the first saccade. It is perhaps only with the advent of the theory above that this effect becomes important, and it has not previously been investigated.

Further, although it was previously known that humans have difficulty distinguishing the two different eyes of origin (Wolfe & Franzel, 1988), complete blindness to eye of origin has not been shown to coexist with an attentional attraction by the very same eye-of-origin feature except in the case of covert attentional cueing without a gaze shift (Zhaoping, 2008a, see below). Indeed, in the case that all the bars had the same veridical luminance contrast against a blank background, and when guided by the experimenter and given sufficient time, most observers could identify the ocular singleton bar as the one appearing brighter or having higher contrast (Zhaoping, 2008a). Further even when the stimulus was shown for only 200 ms, which is too brief a time to allow a saccade, observers could guess more accurately than chance whether an ocular singleton was present. That they might have used the illusory contrast cue is suggested by the observation that they could no longer do this when there was sufficient random variation between the bars in their veridical luminance contrasts. Nevertheless, the cueing effect of the very same ocular singleton, manifested in its effect on the visual search task, did not depend on the uniformity of the contrast of the bars (Zhaoping, 2008a). This suggests that covert attentional cueing by the ocular singleton can be dissociated from the awareness of its presence. What is not clear is whether an overt saccade to this singleton could make it visible to awareness.

Motivated by the questions raised above, the present study uses the rare combination of gaze tracking with dichoptic stimulus presentation to investigate a set of questions. I first tested the prediction that when experimental conditions favor bottom-up influences, a perceptually much less distinctive ocular singleton that is a task-irrelevant non-target can indeed outcompete a more distinctive orientation singleton target in attracting the first saccade in a speeded

visual search task. This prediction is a logical consequence of the definition of saliency, the previous finding suggesting that an ocular singleton is more salient than an orientation singleton, and the V1 saliency hypothesis, to which it would duly lend support.

Second, I assessed whether observers remain unaware of the ocular singleton even when it captures their gaze, if they are unaware of it when it does not (Zhaoping, 2008a). The answer should shed further light on the relationship between action and awareness. If observers become aware of the ocular singleton through gaze capture, the awareness is likely due to a retrospective inference. However, if they remain unaware, we would have identified a dissociation between an action and the awareness of the presence of the sensory input triggering this action that would add to the dissociations seen in previous studies (Belopolsky et al., 2008; Frith et al., 2000; Leach & Carpenter, 2001; Masson et al., 1997; Tavassoli & Ringach, 2010).

Third, I considered how the relationship between the action and the awareness concerned is manifested. For example, I consider whether observers take longer to answer “no” than to answer “yes” to the question of “was there an ocular singleton in the stimulus you just saw?”, if the singleton has captured their gaze.

Findings from two experiments will be reported in this paper. Both experiments used eye tracking during visual search for an orientation singleton target bar among uniformly oriented background bars. [Experiment 1](#) verified that gaze capture by the ocular singleton was indeed the main cause of the cueing effect it exerted in the orientation singleton search reported in the previous study (Zhaoping, 2008a). In order that bottom-up effects should dominate, I took advantage of the facts that visual crowding impairs object recognition but not salient object detection and that crowding is more severe for more eccentric items. Thus, [Experiment 1a](#) used large perceived search images, about  $27^\circ \times 37^\circ$  in visual angle, so that the target was sufficiently eccentric (about  $12^\circ$ ) and crowded by non-targets. Consequently, when the target and an equally eccentric ocular singleton distractor were in opposite lateral halves of the perceived image, guidance to the first search saccade (from the center of the image) was more likely dominated by a competition between the saliencies of the singletons rather than by top-down task-dependent factors due to peripheral target recognition. [Experiment 1a](#) showed that the first search saccade was about three times as likely to be directed to the lateral half that contained the ocular singleton, as to the other half that contained the search target.

However, so that the perceived images could be large, [Experiment 1a](#) used stereo goggles to present dichoptic displays, forcing it to employ the less accurate method of electrooculography (EOG) for gaze tracking. [Experiment 1b](#) used more accurate video gaze tracking but a smaller perceived search image ( $18.7^\circ \times 20.1^\circ$ ) viewed through a four-mirror stereoscope (see [Figure 4A](#)). This thereby verified that the first saccades that were directed to the

lateral half containing the ocular singleton indeed landed close to the singleton distractor. However, the less eccentric target (at 7.3° eccentricity) in **Experiment 1b** was less crowded, enabling a stronger influence of top-down factors on the first search saccade and thereby lessening the gaze capture rate by the ocular singleton distractor.

**Experiment 2** involved assessing subjects' awareness of the ocular singleton when the search stimulus was backward masked after their gaze reached the target. Observers had difficulty reporting after each search trial whether an ocular singleton non-target had been present, even in trials in which their gaze had been captured by the ocular singleton. Preliminary findings of this study have been reported in a conference (Zhaoping, 2010a).

## Materials and methods

In each trial of **Experiment 1**, the search stimulus was randomly presented in monocular (M), dichoptic congruent (DC), or dichoptic incongruent (DI) mode (see **Figure 1**). Observers were asked to press a button as quickly as possible to report whether the target was in the left or right half of the perceived image. In **Experiment 2**, M and DI trials were randomly interleaved, and observers were asked to look at the target (whose position needed to be searched for the purpose) as quickly as they could, whereupon a binocular mask replaced the search display. Observers were then asked to make a non-speeded, forced-choice button press to indicate whether there had been an ocular singleton among the non-target bars. **Experiment 1a** had the virtue of using the same stimulus design and setup as in the previous experiment (Zhaoping, 2008a) and so could verify the cueing effects found before. **Experiments 1b** and **2** used video eye tracking, which is more accurate but was therefore restricted to smaller perceived search images. Further, whereas **Experiment 1a** used bright bars on a black background, **Experiments 1b** and **2** used gray or black bars in a white background to avoid excessive pupil dilation (see **Figures 4A**, **4B**, and **5**).

I assessed differences between real-valued quantities using a two-tailed  $t$  test and between frequencies using a  $\chi$ -square test. Differences were considered to be significant when the  $p$  value was  $p < 0.05$ . Comparisons are accompanied by the  $t$  values or  $\chi$ -square values, degrees of freedom ( $df$ ), and  $p$  values in the  $t$  test or  $\chi$ -square test in the format of  $(t(df), p)$  or  $(\chi^2(df), p)$ , respectively. When an average quantity across subjects for one condition is compared with that for another condition in a statistical test, a matched sample  $t$  test is used. The error bars on the data plotted in all figures denote the standard errors of the mean. Readers not interested in the details of experimental methods may skip directly to the **Results** section.

## Participants

**Experiments 1a**, **1b**, and **2** involved four, three, and eight participants respectively. Each participant was involved in only one of **Experiments 1a**, **1b**, and **2**. All were adults between 18 and 45 years old and had normal or corrected-to-normal vision. They were naive to the purpose of the study, except for one **Experiment 2** participant who could perhaps guess the purpose of the study since he was a vision scientist who had known the previous paper (Zhaoping, 2008a) motivating this study and one in **Experiment 1a** who had known the different dichoptic conditions that could be employed. All demonstrated an ability to see stereo depth, either by correctly judging the relative depth order of two texture layers (as in Zhaoping, 2008a) or the relative depth order of three displayed words and their surrounding frames. Even though the experiments in this study did not involve seeing depth, only subjects demonstrating an ability for depth perception were employed to ensure that the subjects had normal vision in both eyes and that they did not have a lazy eye or some other known or unknown abnormality in monocular vision in any single eye.

## Stimuli

In all experiments, the search target in each visual search stimulus was an orientation singleton bar in a background of uniformly oriented bars in the perceived image. The search display stayed on until after the observers reported the target's location by a button press (in **Experiment 1**) or by gazing at it (**Experiment 2**). In each trial of **Experiment 2**, the visual search stimulus was replaced by a binocular mask stimulus after subjects started gazing at the search target; the mask remained until after the observers pressed a button to report whether they thought there was an ocular singleton non-target in the search display.

Stimuli in **Experiment 1a** were the same as those in the original experiment in which the cueing effect was found (Zhaoping, 2008a), as far as possible within the constraint of the added requirement for gaze tracking. This was to probe gaze behavior for the known cueing effects. The only modifications from the previous stimuli were given as follows: (1) the viewing distance increased from 40 to 50 cm to accommodate EOG recording, scaling down the stimulus size in visual angles accordingly, and (2) the binocular central fixation dot before search started in each trial was flashed (at 2.5 Hz) for one second to encourage central fixation and initialize EOG gaze measurement at the center of the display.

**Figure 1** illustrates the images shown to the two eyes in **Experiment 1a**, together with their superposition, which resembles the perceived image. The images for the left and right eyes were presented in alternate frames on a

Clinton Monoray cathode ray tube (CRT) at 150-Hz frame rate, viewed by the respective eyes with the FE-1 shutter stereo goggles. There was no sensation of image flicker. The full perceived image spanned about  $27 \times 37$  degrees in visual angle and contained 22 rows by 30 columns of identical bars, all uniformly tilted  $\pm 25^\circ$  from horizontal, except for the search target that was tilted  $\mp 25^\circ$  from horizontal and  $50^\circ$  from the other bars. Each bar's position was slightly shifted from the regular texture grid by a random small amount. While each bar was shown to only one eye, vergence was anchored by task-irrelevant binocular dots in between the bars and, throughout each data-taking session, by the CRT frame and four binocular disks ( $0.4^\circ$  diameter) presented at the four corners of the CRT screen. Each trial randomly involved one of the three possible dichoptic presentation conditions (Figure 1): monocular (M), dichoptic congruent (DC), and dichoptic incongruent (DI). In the perceived image, the position for the target or the ocular singleton was always one of the 28 possible row-column locations nearest to a roughly  $12^\circ$  radius circle centered on the fixation, the center of the image where gaze started for each trial. Each of these 28 positions was at least  $9^\circ$  left or right horizontally from the central fixation point. The eye of origin of the target bar, the lateral side of the target, the direction (clockwise or counterclockwise from horizontal) of the tilt of the target bar, and the dichoptic presentation mode (M, DC, or DI) were randomly and independently chosen for each trial. Each trial started with a central (flashing) binocular fixation point on the screen for about one second, then the search stimuli appeared and stayed on the screen until the subject pressed a left or right button to indicate whether the target was in the left or right half, respectively, of the perceived image. The next trial started after another button press by the subject (see Zhaoping, 2008a for more details).

In Experiment 1b, the stimuli and the task were the same as Experiment 1a, except for the design adaptations to use video eye tracking. Video eye tracking is more accurate than EOG, whose spatial resolution is typically no less than  $3^\circ$  in visual angle. The stimuli were made smaller so that the images for the left and right eyes could be displayed side by side on a Mitsubishi 21-inch CRT screen and viewed through a stereoscope with an assembly of mirrors as illustrated in Figure 4A. The bright bars and (the task-irrelevant binocular) dots on a black background in Experiment 1a were transformed into black bars and dots on a white background (with luminance of  $110 \text{ cd/m}^2$ ) in Experiment 1b (Figure 4B), so that the pupil size was not too big. Video eye tracking is inaccurate or impossible if a big pupil is partially occluded by the eyelid. The perceived image had 22 rows by 24 columns of bars (each  $0.6^\circ \times 0.085^\circ$  in visual angle), spanning about  $18.7 \times 20.1$  degrees in visual angle, with an average distance of  $0.85^\circ$  between the centers of the neighboring bars. In each trial, each bar's position deviated randomly from the regular texture grid by up to  $\pm 0.085^\circ$  horizontal and vertically. The targets and ocular singletons were about  $7.3^\circ$  in eccentricity, and at

least  $4.4^\circ$  horizontally, from the center of the perceived image. The central fixation before search stimulus was a static binocular black disk of  $0.21^\circ$  in diameter. All displayed images in each eye, for visual search or fixation, were framed by identical  $19.3 \times 20.7^\circ$  rectangular black frames  $0.17^\circ$  in thickness (see Figure 4B). The search stimulus stayed on screen until 0.3 seconds after the subject's button press response.

In Experiment 2, the visual stimuli were modified from that in Experiment 1b as follows. The orientation singleton target was tilted  $\pm 20^\circ$  from horizontal, and the non-target background bars were all horizontal. Each bar (in the monocular image) was gray, with a random contrast

$$C = 1 - \frac{\text{bar luminance}}{\text{background luminance}}, \quad (1)$$

in the range of  $2/3 \leq C \leq 1$  or  $1/3 \leq C \leq 1$  against the white background (see Figure 5). This non-uniform random contrasts of the bars should make the orientation and ocular singletons less salient than otherwise (when all bars are black with  $C = 1$ ). The display in each trial was randomly in either the monocular (M) or the dichoptic incongruent (DI) condition. The orientation singleton target and the ocular singleton non-target bar were at about  $8.2^\circ$ , and at least  $5.3^\circ$  horizontally, from the center of the perceived image. When the eye tracker verified that the gaze had stayed within  $2.5^\circ$  from the center of the target continuously for half a second, the search target was considered to have been found and the search display was masked by replacing each search stimulus bar with a star shape presented identically to both eyes. Each star shape was made of intersecting bars (of the same shape and size as the search stimulus bars) and had a random luminance contrast  $C$  within the range of  $1/9 \leq C \leq 1$  against the background (Figure 5A). The mask remained present until 0.3 seconds after the subject pressed a button to report whether there had been an ocular singleton among the non-target bars.

## Experimental procedure

In Experiments 1a and 1b, observers were instructed to fixate on the fixation point before the search stimulus appeared but could freely move their eyes to search for the target when the search stimulus appeared. They were asked to press a left or right button as soon as possible to indicate whether the orientation singleton, the search target, was in the left or right half of the perceived image. Observers were not informed that their two eyes would see different images or that each trial could randomly have one of the three possible dichoptic presentation conditions (Figure 1): monocular (M), dichoptic congruent (DC), and dichoptic incongruent (DI). The observers' EOG or video gaze tracking record and their RTs,  $RT_M$ ,  $RT_{DC}$ , and

$RT_{DI}$ , to press the button in the M, DC, and DI trials, respectively, were obtained. Each observer performed 64 search trials for each of the M, DC, and DI conditions; it took about 15 minutes to finish all the search trials.

In **Experiment 1a**, the FE-1 shutter stereo goggles for dichoptic viewing made gaze tracking by video camera difficult; hence, the gaze was tracked using (less accurate) EOG recording (Marg, 1951) at 500-Hz sampling rate. EOG recording involved six EEG electrodes outside the goggle: In each hemi-face, there was one electrode at the temple, one above the eyebrow, and one on the cheekbone (the last two were aligned vertically with the eye). Observers put their chin on a semicircular chin-and-cheek rest, which naturally limited the range of head movements. They were asked to minimize head movements during data taking, although adherence to this instruction was not reinforced. EOG techniques and analysis followed established knowledge (Marg, 1951). EOG was gaze calibrated by asking observers to make a fixed sequence of saccades following a binocular dot moving between the display center and twenty other locations, including ten representative locations for the search target (and the ocular singleton). EOG was also calibrated for blinks and winks, by asking the observers to naturally blink and wink during a calibration period after gaze calibration. These were used to obtain the thresholds for removing the EOG signals evoked by blinks and for detecting winks (none detected during search trials for any observer). This calibration was done for each observer twice before the search trials and repeated twice after to verify the stability of the EOG measurements. The EOG recording was continuous throughout the data-taking session. Times for various stimulus presentations and subject button responses were recorded as additional events to the EOG record in order to synchronize the records of EOG, stimuli, and behavior for data analysis.

In **Experiment 1b**, the gaze from the right eye was tracked through a camera behind a one-way transparent mirror (see **Figure 4A**). This system of stereoscope and video eye tracker was purchased from Cambridge Research Systems ([www.crs Ltd.com](http://www.crs Ltd.com)). Before taking data, eye tracking was calibrated for each observer to within a typical accuracy of 0.5 degree. In each trial, the central binocular fixation was displayed for at least 0.7 seconds, and the search stimulus appeared after the eye tracker verified that the gaze had stayed within  $2.5^\circ$  from the central fixation point for at least 100 ms continuously. Video eye tracking sampled gaze at 250 Hz.

**Experiment 2** used the same equipment setup, tracking calibration and sample rate, and pre-search fixation verification as **Experiment 1b**. Observers had two tasks in each trial, one primary and one secondary. The primary task was to find the orientation singleton target and look at it as soon as possible. When the eye tracker verified that the gaze had stayed within  $2.5^\circ$  from the center of the target continuously for half a second, the primary task was considered complete and the search display was replaced by a mask. The secondary task was to report whether there

had been an ocular singleton among the horizontal bars during the search (see **Figure 5A**). Reports were made with a button press; observers were asked to guess if necessary. Observers were free to move their gaze on the mask and could take their time to press the button. An auditory beep with a low or high pitch sounded after the button press to indicate if the report was correct or incorrect, respectively. Randomly, each trial had either a low or high degree of randomness or variability in the contrasts of the stimulus bars, i.e., the contrast  $C$  of each bar was a random number within a small range of  $2/3 \leq C \leq 1$  or a large range of  $1/3 \leq C \leq 1$ , respectively (see **Figure 5B**). Each subject performed 360 trials, 90 trials for each combination (M-low, M-high, DI-low, DI-high) of dichoptic condition, monocular (M) or dichoptic incongruent (DI), and randomness, low or high, in the bar contrasts. The 360 trials were done in six blocks of 60 trials each, with short breaks in between.

Before data were taken, observers were allowed to examine a DI stimulus at their leisure, with one eye open at a time and with both eyes open. This was so they would know what an ocular singleton was and how it appeared. They were asked to keep both eyes open (except when naturally blinking) during data taking for their tasks and not to slow down their primary task for the sake of the secondary task. They were informed before data taking that the ocular singleton was as likely present as absent in each trial but not informed that the ocular singleton was always on the lateral side opposite to that of the target.

## Other procedures, data analysis, and some notations

In **Experiment 1a**, the EOG signal was notch filtered at 50 Hz and band-pass filtered for the range of 0.3–40 Hz. Horizontal gaze position was fitted (from calibration data, Marg, 1951) to vary linearly with the horizontal EOG, which is the difference between electric potentials of the two temporal electrodes. Similarly, vertical gaze position was fitted to vary linearly with vertical EOG, the difference (averaged across the two eyes) between the potentials from the electrodes above and below the eyes. The time of the first saccade in a search trial was determined as the point that the temporal derivative of the horizontal EOG exceeded a threshold (determined for each subject individually from the EOG for saccades during calibration). The lateral direction of this saccade was determined from the sign of this derivative.

Ignoring the vertical component of the saccadic direction, I denote the horizontal saccadic direction as T-ward (toward the target) if it was in the same horizontal direction as a saccade from the display center to the target. Otherwise, the saccade was reported as B-ward (toward the background). Note that being T-ward or B-ward only denotes the direction, and not the destination, of the saccade. This designation is therefore most meaningful

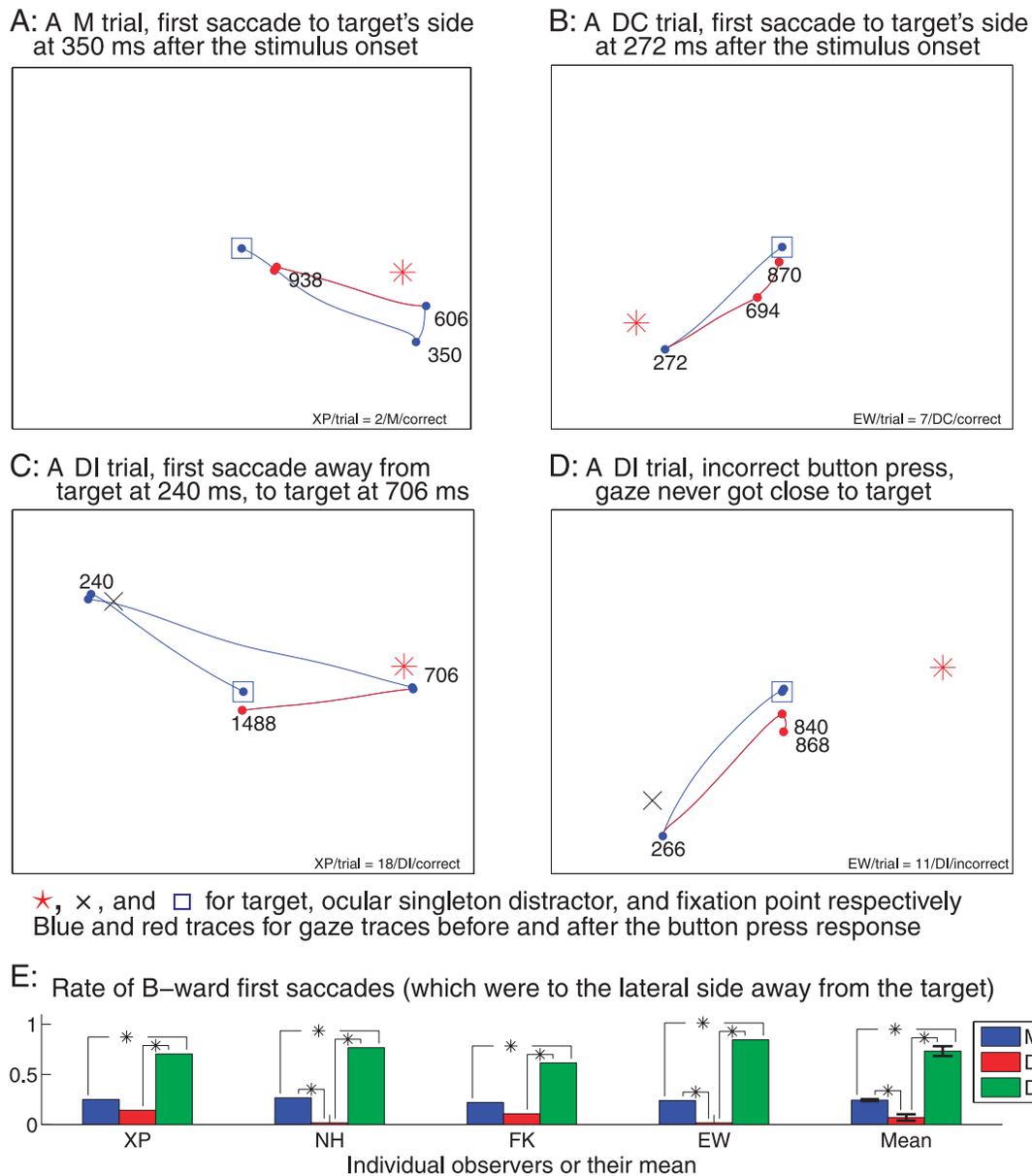


Figure 2. Sample saccades in Experiment 1a. For each dichoptic (M, DC, or DI) condition, and each of the four (XP, NH, FK, and EW) observers, there were 64 trials; 55–64 (mean: 62) of them contained at least one saccade. (A–D) Examples of gaze traces in M, DC, and two DI trials framed by rectangles bounding the perceived image. In each, the blue square indicates the starting position of the gaze; the red star shows the location of the target, and for DI trials, the black star marks the ocular singleton distractor. Gaze traces before and after the button press are in blue and red, respectively. Some numbers near some saccade landings mark landing times (ms) since stimulus onset. (E) Fractions of the first saccades (among all first saccades) that went to the lateral side opposite to the target, i.e., B-ward (background-ward) first saccades (for each observer and averaged across observers). In most of the DI trials, the first saccade went toward the lateral side containing the ocular singleton, away from the target. In all figures of this paper, an “\*” linking two data points (bars) indicates that the difference between them is significant ( $p < 0.05$ ) according to a  $t$  test or  $\chi$ -square test.

for the first saccade in a trial when it reports on the lateral side of the saccade. For example, in Figures 2A and 2B, the first saccade is T-ward; in Figure 2C, the first saccade is B-ward, and the second saccade is T-ward and brings the gaze to the target's lateral side; in Figure 2D, the first saccade is B-ward, and the second saccade is T-ward but the gaze did not reach the target's side. A gaze capture

from fixation by the ocular singleton in a dichoptic incongruent trial would be B-ward, as in Figures 2C and 2D. Of particular interest for this study is whether the first saccade in a trial was T-ward or B-ward, the RT of this saccade, and how both quantities depended on the dichoptic condition M, DC, or DI. A saccade was considered as being associated with a particular trial if

its time of occurrence since the onset of the search display was within the longer of 1.5 seconds or 0.7 seconds plus the RT for the button press. Two (XP and EW) out of the four observers had sufficiently accurate EOG measurements (partly because they had more stable head positions), such that the positions of their saccade destinations in the calibrated saccade sequence could be linearly fitted from EOG data with a mean error smaller than  $4^\circ$ . From these two observers, four examples of their gaze traces during search are shown in [Figures 2A–2D](#). From all observers in [Experiment 1a](#) (in [Figures 2](#) and [3](#)), the times and directions of the first saccade and the first saccade T-ward were obtained, as the calibration data indicated that all significant saccades could have their directions and times adequately obtained from the EOG data even when the accuracy on the exact gaze position was poor.

Calibration of the EOG signal showed that the EOG signals were stable and reliable. In particular, EOG signals were obtained for each calibration sequence, when subjects followed a pre-programmed sequence of dots on the screen. For any single subject, these EOG signals for a calibration sequence before search trials can be reproduced quite well (in EOG amplitudes and timings) in another calibration sequence after the search trials. This reproducibility was also obtained between subjects, taking into account that there is an overall scale difference between the EOG amplitudes for different subjects (due to different sensitivities of the electrodes on different subjects or other factors). Each eyeblink evoked such a huge spike in the vertical EOG that it can be easily taken out of the data by a simple threshold on the temporal derivative of the vertical EOG.

In [Experiments 1b](#) and [2](#), a trial is defined as a bad trial and removed from further analysis if gaze was untracked (often due to blinking or other factors) in more than 10% of the video frames from the onset of the search stimulus to 0.3 seconds after the button press. The RT for the first saccade was determined when gaze shifted at least  $1.5^\circ$  within 20 ms for the first time since stimulus onset. Only RT larger than 100 ms was included for saccade behavior and button press. The gaze was considered to have arrived at the target when it was within  $2.5^\circ$  from the center of the target. In [Experiment 1b](#), each of the three subjects had no more than 15% bad trials (among 64 trials) for each dichoptic condition (range: 3–14.1%, mean: 9.7%). Among the well-tracked trials, at least 94% had a saccade (range: 94.8–100%, mean: 98.6%), and at least 93% (range: 93.1–100%, mean: 97.9%) had gaze arrived at the search target for each observer and each dichoptic condition. In [Experiment 2](#), except for one observer who had about 15% bad trials, all observers had less than 5% bad trials.

Real-time videos of the tracked eye were examined by the author during data taking to watch out for any abnormal behavior, such as subjects closing the tracked eye (and presumably not closing the other eye) in a way that is not blinking. No abnormal behavior was found. Similarly, the

EOG record could clearly pick up any winking behavior, which was done by order in the EOG calibration sequence. However, no winking was found during the visual search trials.

Right after the data-taking session, each observer was asked to comment on the experiment. These conversations were aimed to find out whether the observers noticed the presence of the ocular singleton (e.g., noticing a distracting non-target bar in some trials) in [Experiment 1](#) even though they had not been informed before the data taking, any strategies they used for their tasks, or any unexpected observations.

## Results

All results reported in this paper are averages across observers, except those explicitly stated as being for individual subjects and those in the sample plots of gaze traces of individual subjects. Across observers, the average RTs (or latencies) are the averages of the corresponding mean RTs (or latencies) of individual subjects over trials; the average rates (or frequencies) of events (such as wrong button presses) are the averages of the corresponding rates (or frequencies) for individual subjects.

### Experiment 1: Gaze capture by task-irrelevant ocular singletons in visual search

#### *Experiment 1a: When the singletons were more eccentric from the center of the search display*

In [Experiment 1a](#), to find an orientation singleton target as quickly as possible, an average observer made at least one saccade in 97% of the trials; 99% of these saccading trials had a target-ward (T-ward) saccade, i.e., a saccade in the same horizontal direction as a saccade toward the target from the central fixation. The likely reasons why observers did not simply fixate on the center and rely on peripheral vision alone to do the task are the following. The search array was dense with 660 elements, and the target was sufficiently far from the central fixation point so that due to crowding observers could not easily see the target's orientation without looking at it. [Figures 2A–2D](#) show examples of gaze traces in one monocular (M), one dichoptic congruent (DC), and two dichoptic incongruent (DI) trials. The M and DC examples both show the first saccade T-ward (target-ward). Meanwhile in the two DI examples, the first saccades were B-ward (background-ward), i.e., in the same horizontal direction as one from the center toward the lateral side opposite from the target, and only the second saccades were T-ward. In [Figure 2D](#), the T-ward saccade was apparently not aimed at the target but simply brought the gaze back from the wrong lateral

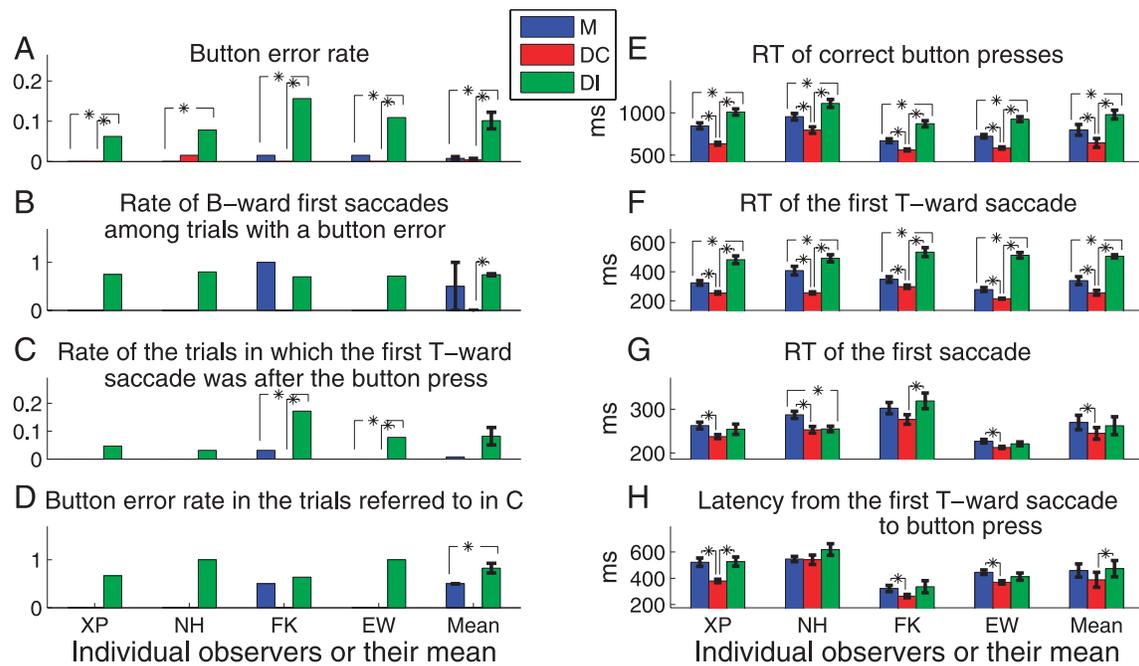


Figure 3. Additional results for each dichoptic condition M, DC, and DI in Experiment 1a for each observer and averaged across observers. (A) Proportions of wrong button responses. (B) Proportions of B-ward first saccades among trials with wrong button responses. (C) Fraction of trials in which the first T-ward saccade (target-ward, i.e., in the same lateral direction as that from central fixation to the target) did not occur until after the button response. (D) Proportion of wrong button responses among trials referred to in (C). (E–H) The average RTs and latencies among trials with a correct button press and T-ward saccades.

side to the display center, while the observer made a wrong button response.

For each of the four observers, 61.4%–84.4% of the first saccades in the DI trials were B-ward, toward the lateral side of the ocular singleton (see Figure 2E). For each observer, the chance for the first saccade to be directed to the lateral side away from the target was significantly higher in the DI than the M and DC trials ( $\chi^2(1) \geq 18$ ,  $p \leq 0.0001$ ). Averaged across the four observers, 73%  $\pm$  5% of the first saccades in DI trials were B-ward, significantly ( $t(3) > 8.7$ ,  $p < 0.004$ ) more than those in the M and DC trials (in which only 24%  $\pm$  1% and 7%  $\pm$  3% first saccades, respectively, went B-ward). This confirms that the lateral side of the non-distinct, task-irrelevant, ocular singleton in the background did regularly outcompete the lateral side of the orientation singleton target to attract gaze for every observer, even though the target was very salient, with its 50° orientation contrast from the background bars. One naive and one non-naive out of the four observers (three of them naive) reported noticing a distracting non-target in some trials.

Figure 3 shows further results from Experiment 1a. Averaged across observers, the button press was incorrect in about 10%  $\pm$  2% of the DI trials. This was significantly ( $t(3) > 4.35$ ,  $p < 0.023$ ) greater than those in the M or DC trials, which had fewer than 1% error trials (Figure 3A). In 74  $\pm$  2% of the DI error trials, the first saccade was B-ward (Figure 3B), suggesting that observers likely

mistook a very salient peripheral object, namely, the ocular singleton distractor, as the search target when they were in too much a hurry and did not adequately verify its identity. Indeed, most of the DI error trials were due to a hurried button press before a T-ward saccade in the DI trials, as in the example of Figure 2D. In 8  $\pm$  3% of the DI trials, i.e., almost as many as the DI error trials, the observers did not make any T-ward saccade before the button press (Figure 3C); 83  $\pm$  10% of such DI trials were error trials (Figure 3D). Hence, of all the DI error trials, 58  $\pm$  8% were those in which the button was pressed before any saccade was made in the direction of the target.

Figures 3E–3H examine the RTs and latencies of various saccades and button responses among trials that had a correct button response and contained T-ward saccades. Note that trials like the example in Figure 2D were not included in these results since they had wrong button presses. First, for each observer, the mean RT for button press was significantly longer in the DI trials than the M and DC trials (Figure 3E;  $t(df) \geq 2.6$ ,  $df \geq 102$ ,  $p \leq 0.01$ ) and was significantly shorter in the DC trials than the M and DI trials ( $t(df) \geq 2.9$ ,  $df \geq 103$ ,  $p \leq 0.004$ ), confirming previous findings (Zhaoping, 2008a). Averaged over observers, the button press RT was 798  $\pm$  64, 644  $\pm$  53, and 979  $\pm$  53 ms for the M, DC, and DI trials, respectively, with a near 200-ms difference between the mean RTs for the M and DI trials. Figure 3F confirms that this result was mainly caused by the influence of dichoptic

condition on the RTs for the first T-ward saccade. Averaged over observers, this RT was  $339 \pm 27$ ,  $255 \pm 17$ , and  $506 \pm 12$  ms, respectively, for the M, DC, and DI trials. For each observer, it was significantly longer in the DI trials than the M and DC trials ( $t(df) \geq 2.15$ ,  $df \geq 102$ ,  $p \leq 0.034$ ) and significantly shorter in the DC trials than the M and DI trials ( $t(df) \geq 2.18$ ,  $df \geq 103$ ,  $p \leq 0.031$ ). This RT pattern is consistent with the observation (Figure 2E) that almost all first saccades in the DC trials were T-ward, while only about three quarters or one quarter of the first saccades in the M or DI trials, respectively, were T-ward. In most DI-trials, the first T-ward saccade was the second or even later saccade during the search.

Meanwhile, the mean RTs of the first saccades regardless of their directions were  $270 \pm 17$ ,  $245 \pm 13$ , and  $262 \pm 21$  ms for the M, DC, and DI conditions, respectively, within 30 ms of each other (Figure 3G). In addition, the mean latencies from the first T-ward saccade to button presses were  $460 \pm 50$ ,  $389 \pm 57$ , and  $474 \pm 63$ , within 90 ms of each other (Figure 3H). Therefore, the latency to initiate the first saccade and that to press the button after a T-ward saccade were not the main causes for slower and faster button responses in the DI and DC trials, respectively.

Interestingly, the average RT (across observers) of the first saccades did not depend significantly on whether it was T-ward or B-ward ( $t(3) \leq 1.757$ ,  $p \geq 0.177$ ), particularly among the DI trials. This suggests that the B-ward first saccades may not be caused by a hurried decision on the saccadic destination, even though the wrong button presses were often caused by a hurried task decision.

Experiment 1a thus confirmed the previous finding (Zhaoping, 2008a) that even though an ocular singleton is not highly distinctive to awareness, it exhibits an effect that is like attention capture by exogenous cueing, shortening or lengthening search button press RT when it is at or far from the target, respectively. Furthermore, Experiment 1a revealed that the initial gaze capture by the (lateral side of the) task-irrelevant ocular singleton was indeed the main cause for slower searches in the DI trials, even though this ocular singleton was far less perceptually distinctive than the target orientation singleton. An orientation singleton with a  $50^\circ$  orientation contrast from background bars was itself very salient, as  $50^\circ$  is several times above the just-noticeable difference for orientation pop out. Apparently, the ocular singleton in the DI trials was even more salient, such that, although it is irrelevant to the task, its lateral side outcompeted that of the orientation singleton target to capture gaze.

Meanwhile, the DC trials had slightly, but significantly, faster first saccades (Figure 3G;  $t(3) = 6.1$ ,  $p = 0.009$ ) and fewer B-ward first saccades (Figure 2E;  $t(3) = 4.7$ ,  $p = 0.018$ ) than the M trials. This suggests that, when there was no ocular singleton distractors (among the background bars), the target captured gaze more strongly when it was also the ocular singleton than otherwise.

Furthermore, the button press followed the T-ward saccade more promptly in the DC trials (Figure 3H) than

in the DI trials (significantly  $t(3) = 3.85$ ,  $p = 0.03$ ) and the M trials (marginally  $t(3) = 2.5$ ,  $p = 0.088$ ). This suggests that the target's being more salient in the DC trials (due to its being an ocular singleton) speeded up the task decision, perhaps by making gaze land closer to the target after the T-ward saccade and/or a more confident decision making after the gaze reached the search target.

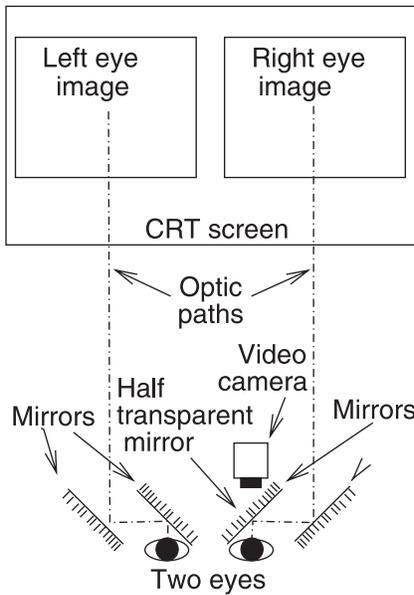
### Experiment 1b: When the singletons were less eccentric from the center of the search display

Since EOG tracking cannot resolve gaze position more accurately than about  $3^\circ$ , Experiment 1a could only distinguish whether the first saccade was toward the lateral half of the search image containing the target or the ocular singleton distractor in the DI trials and could not pinpoint the target or the ocular singleton as the source of the gaze attraction. Experiment 1b used the more accurate method of video eye tracking to verify that the ocular singleton was the saccadic target, albeit at the expense of using a stimulus subtending a smaller range in visual angle.

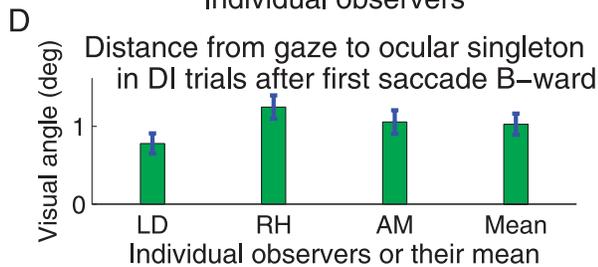
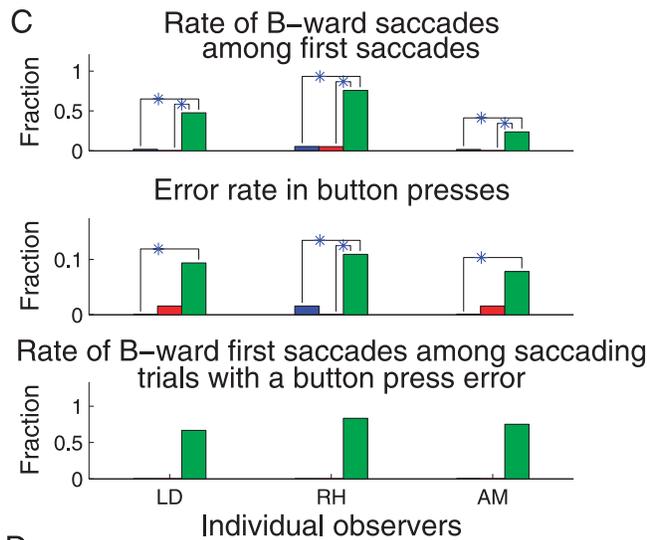
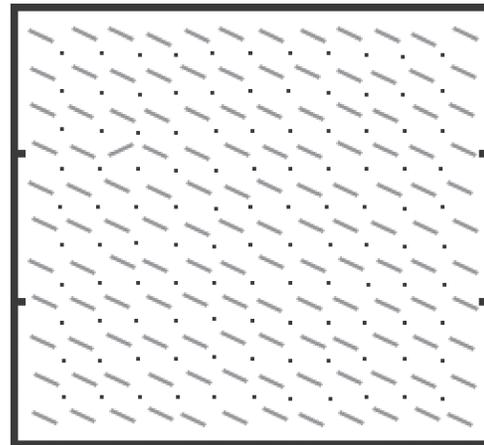
One of the consequences of using a smaller stimulus is that the search target is closer to the display center, making it less susceptible to visual crowding by the non-target (Levi, 2008) and thus more easily identified by observers fixating on the display center at the start of a search trial. Hence, the influence of top-down task-driven factors on the first saccade should be stronger, thereby relatively weakening the influence of bottom-up saliency that could guide a saccade before the saccadic target was identified. Indeed, in each dichoptic condition, observers made fewer B-ward first saccades in Experiment 1b than in Experiment 1a. Nevertheless, bottom-up saliency was sufficiently strong as to reveal gaze capture by the ocular singleton.

All subjects made saccades in more than 94% of the well-tracked trials. In each observer, the rate of B-ward first saccades was significantly larger in the DI than in the M and DC conditions ( $\chi^2(1) \geq 12.7$ ,  $p < 0.0004$ ; Figure 4C). Averaged over the three new observers,  $49 \pm 15\%$  of the first saccades were B-ward in the DI trials whereas only  $3 \pm 1\%$  and  $1.7 \pm 2\%$  of the first saccades were B-ward in the M and DC trials, respectively. In each observer, the button press error rate was significantly larger in the DI trials than the M trials ( $\chi^2(1) \geq 4.8$ ,  $p \leq 0.029$ ), and 67–83% of the DI error trials had the first saccade B-ward (Figure 4C). These findings are consistent with those in Experiment 1a. Hence, the RTs for the first saccade, target arrival, and button press depended on the dichoptic condition in the manner predicted from the attentional and gaze attraction by the ocular singleton (Figure 4E). Furthermore, with its more accurate gaze tracking, Experiment 1b verified that, averaged across observers, the gaze was brought to a mean distance of  $1.03 \pm 0.14^\circ$  from the ocular singleton in the DI trials by the first saccade when it went B-ward (Figure 4D). Meanwhile, the RTs for

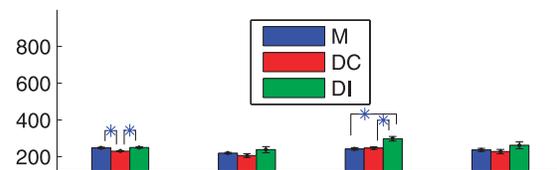
**A** Stereoscope video eye tracking set up



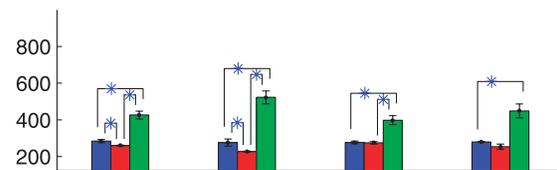
**B** A smaller version of an example perceived search image



**E** RT of first saccades (among saccading trials)



RT of target arrival (among target arrival trials)



RT of correct button presses

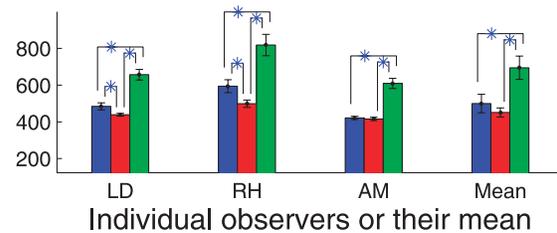


Figure 4. **Experiment 1b** was an adaptation of **Experiment 1a** to use the more accurate method of video eye tracking. For each of the three observers (LD, RH, and AM), and each dichoptic condition (M, DC, or DI), 64 trials were performed; 86–97% (mean: 90.3%) of them were sufficiently well gaze tracked, of which 94.8–100% (mean: 98.6%) contained at least one saccade and 93.1–100% (mean: 97.9%) had gaze arriving at the target. (A) Schematic of the equipment setup. (B) Illustration of the perceived image of a (reduced-size version of a) stimulus example. (C) For each dichoptic condition and observer, the proportion of B-ward saccades among first saccades, the button error rate (regardless of saccades), and the proportion of trials that had a B-ward first saccade among trials with a button error and at least one saccade. (D) Mean distance (for each observer and averaged across observers) between the ocular singleton and the gaze position as the result of the B-ward first saccades in the DI trials. In each DI trial having a B-ward first saccade, this distance is calculated as the shortest distance between the gaze and the ocular singleton within 50 ms from the start of the saccade. (E) RTs for the first saccade, target arrival, and correct button press in each dichoptic condition, for each subject and averaged across the subjects.

the gaze to arrive at target in the M, DC, and DI trials were  $279 \pm 3$ ,  $253 \pm 14$ , and  $449 \pm 38$  ms, respectively (see Figure 4E). In each subject, the DI trials required significantly longer RTs than the DC and M trials for the gaze to reach target ( $t(df) \geq 4.9$ ,  $df \geq 106$ ,  $p < 0.0001$ ) or for the button presses ( $t(df) \geq 3.35$ ,  $df \geq 118$ ,  $p \leq 0.001$ ).

Two out of three (all naive) observers reported noticing the presence of a distracting non-target in some trials.

Experiment 1b also established a background finding for Experiment 2, which used the same stimulus display setup and video tracking method.

## Experiment 2, dual task: Saccade to the orientation singleton and report on whether an ocular singleton was present

Trials in Experiment 2 were randomly picked from M or DI dichoptic conditions. Observers had to complete two tasks in each trial. The primary task was to find the orientation singleton target among the background horizontal bars by shifting gaze to the target as soon as possible. Once gaze had stayed on the target for half a second, the stimulus was then replaced by a binocular mask. The second task was to answer by a forced-choice button press whether there had been an ocular singleton in the search display. Observers were informed before the data taking that an ocular singleton was as likely to be present as absent in each trial.

The main purpose of this experiment was the second task's assessment of the observers' degree of awareness of the ocular singleton, which could have captured attention or gaze during the primary task. The previous study (Zhaoping, 2008a) showed that the performance in a forced-choice test regarding the presence or absence of an ocular singleton reached chance level when the stimulus was presented too briefly to evoke overt saccades during its presentation and when the bar luminance was sufficiently non-uniform or randomized between bars. This held even though the same brief ocular singleton significantly cued attention covertly so that it affected visual performance in an orientation singleton discrimination task and even though this cueing effect was no weaker than when forced-choice answers were made more accurate by dint of making the luminance of the bars uniform.

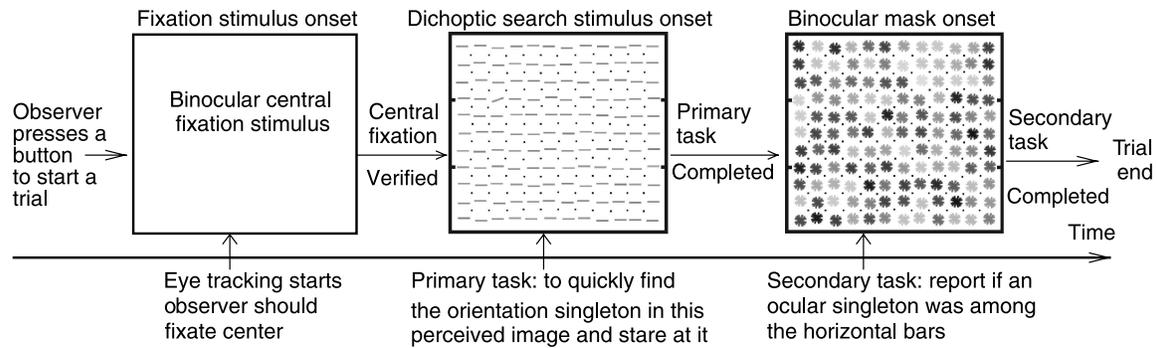
Experiment 2 asked if the ocular singleton could still evade awareness when the stimulus was presented long enough for the ocular singleton to capture overt gaze, and, if not, what would be the relationship between awareness and gaze capture. Each trial had either a low or high degree of randomness in the luminance of the bars, with the contrast ( $C = 1 - (\text{bar luminance}/\text{background luminance})$ ) of each bar against background being a random number in the range of  $2/3 \leq C \leq 1$  or  $1/3 \leq C \leq 1$ , respectively (Figure 5B).

Figure 6 shows some examples of gaze and button response behavior for a typical subject in Experiment 2. The gaze behavior during the search for the orientation singleton was qualitatively as expected from the findings in Experiment 1. Meanwhile, even when the gaze was captured by the ocular singleton during the search (Figure 6E), observers could report that no ocular singleton was presented. Figure 6B shows an example in which the observer reported ocular singleton as being present even though it was not and even though the gaze went directly to the target during the search.

Averaged across eight observers, significantly more first saccades were B-ward in the DI than the M trials (for either randomness condition,  $t(7) \geq 4.1$ ,  $p \leq 0.0045$ ) during the search phase before the mask onset (see Figure 7A). The rates of B-ward first saccades were about 13% in the M trials regardless of the randomness condition, but were  $37 \pm 6\%$  and  $27 \pm 5\%$  in the DI trials in low and high randomness conditions, respectively. Gaze was captured in  $31 \pm 6\%$  and  $17 \pm 3\%$  of the DI trials in low and high randomness conditions, respectively (Figure 7B), some of them as the result of the first saccades, others by subsequent saccades. As a consequence, RT for gaze to reach target during search was significantly ( $t(7) \geq 2.85$ ,  $p \leq 0.025$ ) shorter in the M than the DI conditions (Figure 7C). In the DI trials, the rates of B-ward first saccades were lower than those in Experiment 1b, mainly because background inhomogeneity (by the randomness in the bar contrasts) made the ocular singletons less salient. For the same reason, and additionally due to the smaller orientation contrast ( $20^\circ$ ) of the target, the orientation singleton target was also less salient than it was in Experiment 1b, making the rate of B-ward first saccades in the M trials higher than that in Experiment 1b. This was manifested in the longer RT (400 ms or longer) for the gaze to arrive at the target in the M trials (often not by the first search saccade, which had an average RT around 250 ms).

In the secondary task of reporting whether there was an ocular singleton among the horizontal bars, the miss rates in the DI trials were  $0.26 \pm 0.06$  in the low randomness condition and  $0.5 \pm 0.03$  (equivalent to that of unbiased pure guessing) in the high randomness condition. Even among the DI trials in which the gaze was captured by the ocular singleton, the miss rates were  $0.21 \pm 0.09$  and  $0.37 \pm 0.05$  for the low and high randomness conditions, respectively. As each pure and unbiased guess has a 50% error rate, these results suggest that, even when the ocular singleton captured gaze, the observers were so uncertain that their reports resembled those of somebody who was purely guessing at the presence of an ocular singleton in two fifths or two thirds of the times (to get an error rate of one fifth or one third), in the low or high randomness condition, respectively. However, observers were more certain about the absence of the ocular singleton among the M trials. Even in the high randomness condition, the correct rejection rate was  $69 \pm 5\%$  in the M trials, significantly

## A: The time sequence of events within a trial in exp. 2



## B: Example perceived search stimuli for the primary task in exp. 2 with low (left) and high (right) degrees of randomness in the contrasts of the search stimulus bars

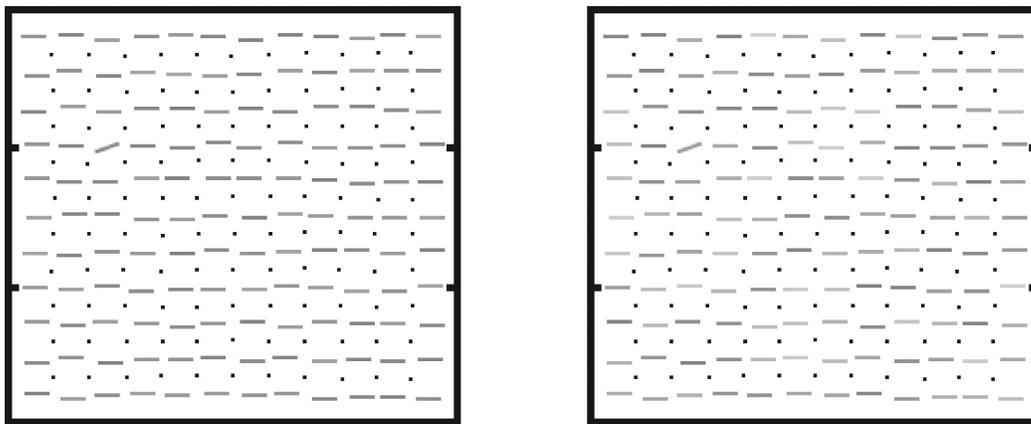


Figure 5. **Experiment 2** to study the awareness of the ocular singleton in gaze captures. (A) Design of **Experiment 2** illustrated in the time sequence of events in a trial having a primary and a secondary task. The illustrated search and mask stimuli are smaller versions of a random perceived search image and a random mask. (B) Smaller versions of random perceived stimuli for the low and high randomness conditions when the bar contrasts against the background had, respectively, low and high variability between the bars. Trials involving the low and high randomness were randomly interleaved.

larger than the miss rate of  $50 \pm 3\%$  in the DI trials ( $t(7) = 3.14$ ,  $p = 0.016$ ), even when excluding the DI trials in which the ocular singleton captured the gaze ( $53 \pm 3\%$ ;  $t(7) = 2.7$ ,  $p = 0.031$ ).

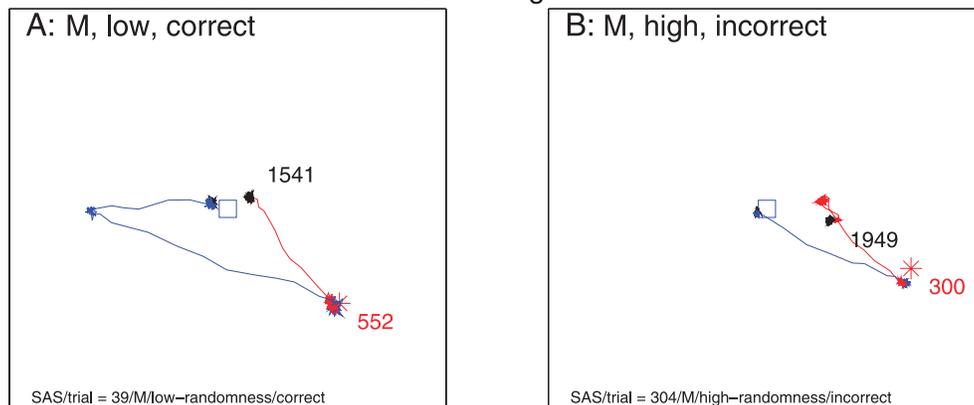
**Figure 8** further examines the relationship between the awareness of and gaze capture by the OS. When the randomness among the bar luminance was low, the miss rate did not significantly ( $t(7) = 0.84$ ,  $p = 0.43$ ; **Figures 8A** and **8B**) depend on whether the ocular singleton captured gaze. The opposite also holds, i.e., the rate of gaze capture by the ocular singleton did not significantly depend on whether the ocular singleton was reported as present ( $t(7) = 0.72$ ,  $p = 0.50$ ). Thus, awareness of the ocular singleton was dissociable from gaze capture by this singleton.

However, this state of affairs changed when the randomness was high. In this case, the hit rate was significantly higher if the ocular singleton captured gaze than otherwise ( $t(7) = 3.12$ ,  $p = 0.017$ ), and the gaze capture rate was

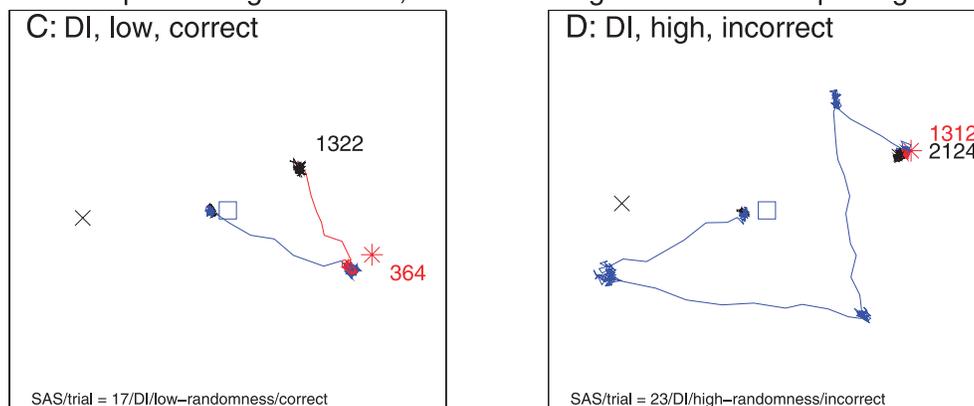
significantly higher among the ocular singletons reported as being present than that among the other ocular singletons ( $t(7) = 2.9$ ,  $p = 0.023$ ; **Figures 8A** and **8B**). According to typical reports from observers and the author's own experience, an ocular singleton is more recognizable when the visual items in the display are sufficiently uniform in veridical contrast, as the ocular singleton typically appears to have a different (typically stronger) contrast, which can be used as a cue to identify the ocular singleton. Hence, **Figures 8A** and **8B** suggest that observers might link gaze capture with the presence of an ocular singleton only when they could not exploit this "illusory" contrast cue for ocular singleton recognition, as this cue was more submerged within the contrast variance in the high randomness condition.

Meanwhile, although every observer was asked for their observations and comments immediately after data taking, only one out of eight observers reported noticing a weak

## Monocular trials without an ocular singleton



## Dichoptic incongruent trials, the ocular singletons did not capture gaze



## Dichoptic incongruent trials, the ocular singletons captured gaze

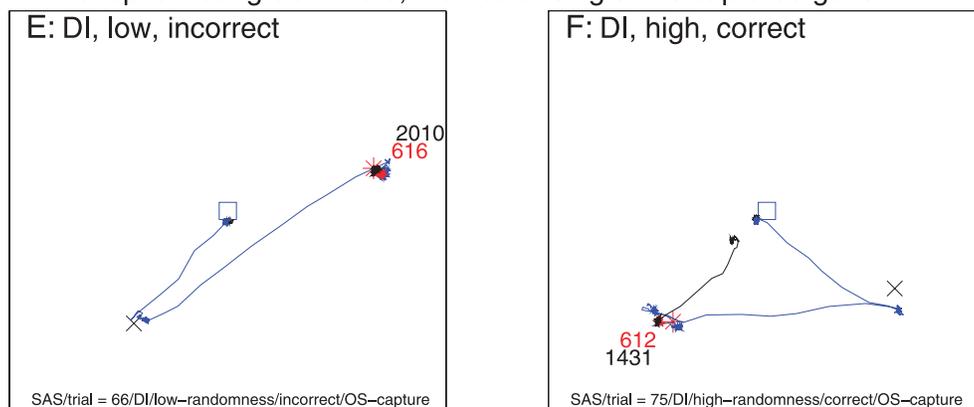


Figure 6. Some gaze traces of a typical subject, framed by rectangles indicating the boundaries of the perceived search images in Experiment 2. Plotting conventions are as in Figures 2A–2D except that the blue traces are during the target search before the mask onset (which occurred once gaze had stayed continuously with the target for 0.5 seconds), the red traces are between mask onset and the button press reporting the presence or absence of the ocular singleton (OS), and the black traces are after the button press. Numbers in red and black, respectively, indicate the RTs in ms (since search stimulus onset) of gaze arrival to target (and then staying for at least 0.5 seconds) and that of the button press. Dichoptic condition (monocular (M) or dichoptic congruent (DI)), degree (low or high) of randomness in the luminance of the bars, and button press accuracy (correct or incorrect) are marked in each example. Gaze is considered captured by ocular singleton in (E) and (F) but not in (D). First saccade went B-ward in (A), (D), (E), and (F).

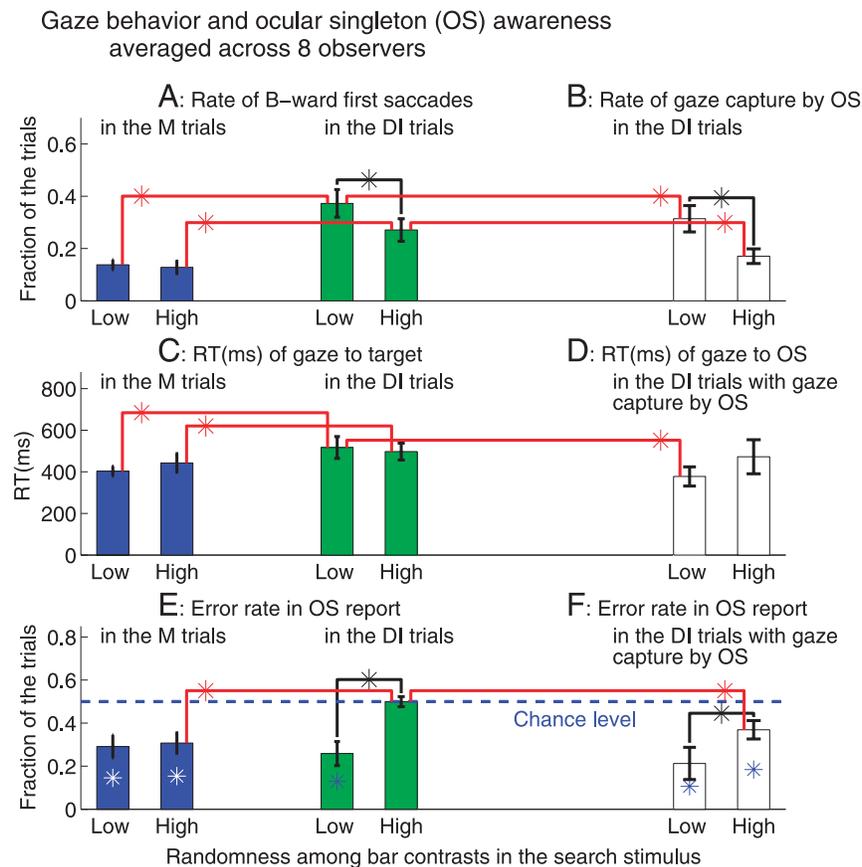


Figure 7. Gaze behavior and the awareness of the ocular singleton (OS) in Experiment 2 averaged over 8 observers. Each observer performed 90 trials in each combination of dichoptic condition (M, when an ocular singleton was absent, or DI, when an ocular singleton was among the non-target bars) and randomness condition (low or high degree of variability in the bar contrasts). (A) Rates of B-ward first saccades in the M or DI trials and (B) rates of gaze capture by the ocular singleton in the DI trials, during the primary task. (C) RTs during the primary task for gaze to reach an orientation target in the M or DI trials or (D) to reach ocular singleton in the DI trials in which gaze was captured by ocular singleton. (E) Error rates in reporting whether an ocular singleton had been present during the primary task in the M or DI trials or (F) in those DI trials in which the gaze was captured by the ocular singleton. In (E) and (F), an “\*” on the data bar indicates that it is significantly different ( $t(7) \geq 2.7$ ,  $p \leq 0.03$ ) from the chance level of 0.5.

(and unreliable) link between gaze wandering during the search and the presence of an ocular singleton (indicated to the observers by a feedback to their report after each trial). This suggests that most observers might only subconsciously associate the presence of the ocular singleton with gaze wandering behavior. If such an association was present, RTs for the primary visual search should be longer (to allow gaze wandering) among trials reporting ocular singleton as present in the secondary task. Figure 8C shows that this holds significantly among the high randomness trials ( $t(7) = 2.69$ ,  $p = 0.031$ ) or when considering all trials regardless of the randomness ( $t(7) = 2.71$ ,  $p = 0.030$ ). Whether the observers reported an ocular singleton was present affected the average search RT by only about 60 ms, much shorter than a typical inter-saccadic interval. This suggests that the association between the gaze capture by the ocular singleton and the awareness of the ocular singleton was quite weak.

Observers took much longer (following mask onset) to report ocular singleton as absent in the DI trials if their gaze was captured by the ocular singleton, as if they were hesitating about their erroneous report (Figure 8D). If the ocular singleton captured gaze, the average RT to answer “absent” was significantly longer than that to answer “present” ( $t(7) = 2.77$ ,  $p = 0.028$ ), by about 600 ms. In addition, the average RT to report ocular singleton as absent in the DI trials was significantly longer if the gaze had been captured than otherwise ( $t(7) = 3.1$ ,  $p = 0.017$ ). Meanwhile, the RTs for these erroneous “absent” reports in gaze capture trials were marginally but not significantly ( $t(7) = 2.25$ ,  $p = 0.059$ ) different from those for the erroneous “present” reports in the monocular (M) trials—both were incorrect reports in the secondary task, one missing the ocular singleton despite the gaze capture and the other false alarms. There was no significant difference ( $t(7) \leq 1.46$ ,  $p \geq 0.19$ ) between the RTs for the

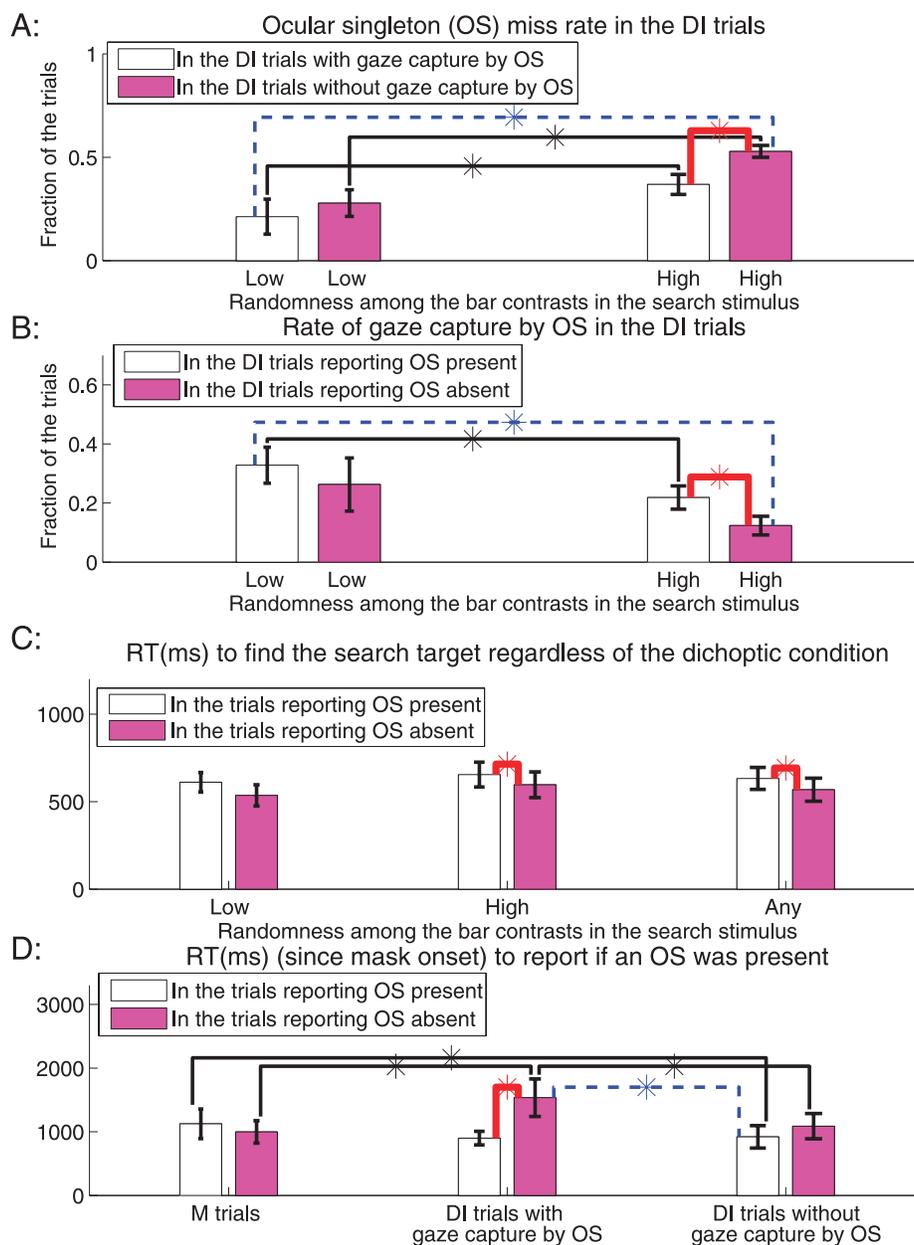


Figure 8. Awareness of the ocular singleton (OS) non-target and its propensity to capture gazes in Experiment 2 (averaged across 8 observers). (A) Miss rates among the DI trials with and without a gaze capture by the ocular singleton. (B) Rates of gaze capture by the ocular singleton among the DI trials in which the ocular singleton was reported as present and absent, respectively. (C) RTs to find the orientation singleton target (when the gaze started waiting at the target for mask onset) among trials having a subsequent report in the secondary task that an ocular singleton was present or absent, respectively. (D) RTs (since mask onset) to report if an ocular singleton was present and absent, respectively, among the monocular (M) trials, the DI trials with the gaze capture by the ocular singleton, or the DI trials without gaze capture by the ocular singleton, respectively, regardless of the randomness conditions.

three types of correct reports: hits with gaze capture, hits without the gaze capture, and correct rejections. Averaged across dichoptic conditions, the RTs for the incorrect reports,  $1143 \pm 213$  ms,  $1098 \pm 189$  ms, and  $1180 \pm 230$  ms, respectively for all, low, and high randomness conditions, were slightly but significantly ( $t(7) \geq 2.53$ ,  $p \leq 0.039$ ) longer than the corresponding RTs ( $983 \pm 168$  ms,  $932 \pm$

$143$  ms, and  $1046 \pm 200$  ms, respectively) for the correct reports.

To summarize, Experiment 2 found the following: When the search items had sufficiently similar luminance contrasts, observers were significantly aware of the ocular singleton, perhaps through its illusory contrast cue. Awareness did not depend on whether the singleton captured their

gaze. When the search items were sufficiently dissimilar in luminance contrasts, observers' awareness was much reduced but was significantly and positively associated with the gaze capture. Their residual awareness was partly manifested in the dependence of their RT for reporting the ocular singleton, on whether the report was correct and whether their gaze was captured.

## Discussion

### A stronger attentional capture by a perceptually less distinctive input

One of my main observations is that a singleton in eye of origin, a visual feature that is nearly invisible to perception, can capture gaze more strongly than a much more perceptually distinctive orientation singleton. This finding is distinguished from previous findings of attentional attraction by stimuli that are perceptually hard to distinguish or detect (Jiang et al., 2006; McCormick, 1997; Mulckhuyse, Talsma, & Theeuwes, 2007) because the usual relationship between attentional attraction and perceptual distinctiveness is reversed and because of gaze capture. In the previous studies, saliency always increases with the degree of perceptual distinctiveness in the feature contrast used for cueing. Hence, their perceptual invisibility or indistinction had to be achieved by reducing the cueing feature contrast (e.g., using a dark gray bar on a black background in McCormick, 1997), or making the cue be brief (e.g., only 15–16 ms in McCormick, 1997 and Mulckhuyse et al., 2007), or by very salient masks (e.g., a static image in one eye suppressed by dynamic random noise in the other eye in Jiang et al., 2006). Consequently, the cueing effects found in the previous studies were never so strong as to capture gaze but were merely manifested in shortened RTs or increased performance accuracies of other tasks. They were also only significant when averaged over multiple subjects (Jiang et al., 2006; McCormick, 1997; Mulckhuyse et al., 2007). In contrast, the cueing feature contrast in the current study is in eye of origin, which is naturally obscure to perception since it is not explicitly represented in higher brain areas. For ocular contrast, there is no tight link between perceptual distinctiveness and saliency. Thus, the strength of the cueing ocular contrast can be maintained while the cueing item is made perceptually obscure (by camouflaging the illusory contrast of the ocular singleton through inhomogeneous luminance). Consequently, the cueing effect in the current study is strong enough to capture gaze, sometimes even more strongly than a perceptually more distinctive target feature contrast, and is significant even in individual subjects.

It should be noted that the attentional attraction of the ocular singleton cannot be accounted for by a singleton search mode, i.e., the observers searching for just unique or

odd feature value. This is because the cueing effect by the ocular singleton was very strong and significant (shortening RT by more than half) even when the search was for a letter “T” among letter “L”s (Zhaoping, 2008a), when observers were not searching for a unique orientation singleton.

### Awareness of an ocular singleton through its illusory contrast cue and gaze attraction

Reports from observers suggest that when input luminance (or contrast against background) of various visual input items are uniform, an eye-of-origin singleton could appear to have a distinct contrast from other items. Indeed, in [Experiment 1](#) when all bars had the same veridical contrast, half of the naive observers spontaneously reported noticing a distracting non-target in some trials (when asked after data taking for their comments on the experiment).

In [Experiment 2](#), among trials in which bar luminance were less variable between stimulus items, the illusory contrast cue likely enabled observers to miss only  $26 \pm 6\%$  of the ocular singletons in the secondary task. Remarkably, this miss rate did not depend on whether the ocular singleton captured the gaze, suggesting that, when observers could exploit the illusory perceptual cues, their awareness of the ocular singleton could be dissociable from their action of gaze capture by the very same singleton.

However, this dissociation did not hold when the input contrasts were sufficiently non-uniform—the miss rate rose to  $53 \pm 3\%$  among the ocular singletons that failed to capture gaze but only to  $37 \pm 5\%$  among those that did capture gaze, indicating some degree of awareness (Shanks, 2010) through gaze capture. Two previous studies (Wolfe & Franzel, 1988; Zhaoping, 2008a) have indicated that, when the input luminance are sufficiently non-uniform, human observers become completely unaware of the presence of the ocular singleton. In one of these two studies (Zhaoping, 2008a), the stimulus was too brief to allow saccades during its presentation; in the other study (Wolfe & Franzel, 1988), the stimulus was presented long enough (for the extremely frustrating task of finding an ocular singleton target) to allow multiple saccades, making it hard for gaze capture to appear special. These observations suggest that, when the illusory contrast cue is largely submerged by non-uniform input contrasts, the limited awareness of the ocular singleton is unlikely to be due to any perceptual visibility of the eye-of-origin feature but rather to an association between an evoked saccade to non-targets and the presence of this feature (which was indicated to my observers by the feedback they received on every trial). My data suggest that this association is quite weak.

### Action, perception, and awareness

My observers often reported that the ocular singleton was absent even when they actively looked at it. This is a

case of action without perception (Georgeson, 1997) and can be seen as both conventional and remarkable. Humans can often be unaware of some components of their actions. For example, they may not be aware that their hand trajectory deviated when attempting to grasp an object that suddenly shifted in position (Frith et al., 2000) or of the deviations in their saccadic trajectory toward a distinct distractor that suddenly appears on the scene when they are saccading toward a target (Belopolsky et al., 2008). Hence, if humans were unaware of the ocular feature to start with, it is not surprising that they could not be made aware by a saccade caused by this feature. This could happen if observers did not associate the saccadic action with the feature or if the saccadic action itself also evaded awareness. There have also been other observations that perception does not fully benefit from the higher accuracy of our motor system. For example, humans prefer to saccade to the slightly earlier one of the two onset targets even when their perception cannot distinguish which one is earlier (Leach & Carpenter, 2001), and their gaze or vergence pursuits can follow a change in velocity or the direction of a disparity shift of visual stimuli more accurately than their perception (Masson et al., 1997; Tavassoli & Ringach, 2010).

However, it is critical to note that normal humans are aware of the presence of the object of their grasp, of the chosen saccadic target among the two target choices, of their gaze pursuit target, and of the occurrence of a sudden change (in a visual input disparity) that drives their vergence pursuits (Masson, private communication, 2010). It is also most likely that humans would have been able to report the sudden appearance of the distinct distractor (in the task of Belopolsky et al., 2008) if they had been asked to do so as a secondary task (with the primary task being to saccade to a target). It is thus remarkable, and hitherto unreported, that my observers had difficulty with their task of reporting the presence of the ocular singleton that was nevertheless sufficiently salient to capture an overt saccadic action.

The observation that gaze capture itself can cause no more than a limited awareness of (the presence of) the visual feature attracting the gaze is consistent with observations from two previous studies (Zhaoping & Frith, 2011; Zhaoping & Guyader, 2007). In these two studies, the gaze capturing salient feature was in fact a naturally visible orientation singleton bar oriented at least 45° away from the non-target bars and was the target of a visual search task but was camouflaged perceptually. In Zhaoping and Guyader (2007), for example, the target was an oblique bar tilted 45° from vertical among non-target bars tilted 45° from vertical in the other direction. However, each of these bars was intersected by a task-irrelevant horizontal or vertical bar, forming an “X” shape like “”, “”, “” or “”, each was a rotated version of another. Consequently, the target bar was camouflaged, since, due to rotational invariance of shape recognition, the “X” shape containing the target bar was identified

with, and confused by, the neighboring “X” shapes made of non-target bars. Although the unique tilt of the target oblique bar was salient enough (among hundreds of non-target items) to attract attention and gaze, the gaze shift to the target was often insufficient to make observers aware of the target bar, such that the gaze often subsequently abandoned the target to search elsewhere.

Meanwhile, findings from another study (Zhaoping, 2008b) suggest a seemingly stronger link between gaze shifts and performance in task decision, although observers were apparently unaware of such a link. In about half of the trials of a visual search for an orientation singleton target among hundreds of non-target items, the search stimulus was masked before observers’ gaze reached the target—these trials are called after-search trials. Observers had to report in each trial, without time pressure and guess if necessary, whether the target was in the left or right half of the display. Among the after-search trials, observers made on average  $3.1 \pm 2.8$  saccades on the mask before reporting. If their gaze reached the location of the vanished target (as happened in a minority of these after-search trials), their report had about an 84% chance to be correct, regardless of whether it was the first or a subsequent saccade (after the mask onset) that brought the gaze to the target. On average, these correct reports occurred about half a second after the gaze arrived at the location of the vanished target, and a small fraction of these correct reports occurred immediately before the gaze arrival as if in anticipation. In contrast, if gaze never reached the location of the vanished target during the after-search, the accuracy of the reports was at the chance level (50%). After the experiment, observers reported that they had to guess for their reports in the trials in which they did not see the target, and none of them reported any link between their guesses and their gaze shifts. The relationship between the findings in this previous study and those in the current study is not clear. It is possible that a shorter latency between the gaze arrival and the observers’ report may have helped the decision performance in the previous study.

## The neural substrate for gaze attraction by bottom-up saliency

Since explicit information about the eye of origin of visual input is barely available beyond V1 along the visual pathway (Burkhalter & Van Essen, 1986; Hubel & Wiesel, 1968; Zeki, 1978), the current finding supports the theory (Li, 1999a, 2002) that a bottom-up saliency map is created in V1. This area is lower in the visual hierarchy than traditionally thought of for a feature-independent quantity such as saliency. However, evidence in favor of the theory is accumulating. For instance, recent observations using functional magnetic resonance imaging and EEG event-related potentials suggest that V1, rather than the frontal and parietal cortical regions more closely associated with

perceptual awareness, is activated by a salient orientation singleton that eluded perception (due to masking) while significantly cueing a visual discrimination task at its location (Zhang, Zhaoping, Zhou, & Fang, 2012).

The weak association I found between gaze capture and the awareness of the stimulus that led to this motor action suggests that V1 may cause such gaze shifts directly through its monosynaptic connection to the subcortical superior colliculus, which controls gaze, rather than indirectly via the higher brain areas, like the frontal eye field. The latter pathway may be more important to execute for task-dependent eye movement intentions (Hasegawa, Peterson, & Goldberg, 2004; Schiller, 1998; Tehovnik, Slocum, & Schiller, 2003).

The main intra-cortical mechanism in V1 responsible for saliency computations is iso-feature suppression (Li, 1999a). This makes V1 neurons tuned to similar visual features, such as orientation, color, motion direction, or eye of origin, suppress each other (Allman, Miezin, & McGuinness, 1985; C. Y. Li & Li, 1994). These mechanisms have previously been observed to cause the contextual influences observed physiologically, e.g., the typical suppression of the activity of a neuron coming from stimuli outside its classical receptive field. In my stimuli, the background bars are subject to iso-orientation suppression from nearby and identically oriented neighboring bars (Knierim & Van Essen, 1992; Sillito, Grieve, Jones, Cudeiro, & Davis, 1995), and bars presented to one eye are subject to iso-ocular suppression from nearby bars presented to the same eye (DeAngelis, Freeman, & Ohzawa, 1994). The orientation and ocular singletons both escape this iso-feature suppression, allowing them to attract attention. This process has been demonstrated in a model of V1 (Li, 1999a, 1999b). The intra-cortical interactions and the contextual influences motivated the computational framework of segmentation without classification (Li, 1999b), which involves segmenting one image region from another (by highlighting V1 responses to region boundaries) without classifying or recognizing the regions first. This framework is manifested particularly in relatively higher V1 responses to salient feature singletons, regardless of what visual features cause the saliency. Evidence for this also comes from texture segmentation and visual search behavior involving orientation, color, and motion features supporting this framework (Koene & Zhaoping, 2007; Zhaoping & May, 2007).

## The role of ocular dimension in visual attention and perception

My results also suggest that ocularity is a potent basic visual feature dimension for bottom-up attentional attraction (Zhaoping, 2010b). Indeed, I found it to be just as strong as, or even stronger than, orientation, which is the paradigmatic basic feature dimension (Treisman & Gelade, 1980). Ocularity has so far been overlooked as a feature

dimension due to its limited access to awareness, making impossible a search requiring a conscious report of an ocular singleton (Wolfe & Franzel, 1988). Recognizing ocularity as a basic feature dimension requires realizing that the strength of attentional attraction does not necessarily increase with the perceptual distinctiveness of the sensory input causing this attraction.

Why might ocularity lead to gaze capture even without leading to substantial perceptual awareness? In one sense, this is the wrong question: If vision is looking and perceiving, the first looking should perfectly well occur without perceiving but serving to facilitate subsequent perceiving. For example, gaze shift in response to ocular contrast signals may facilitate figure-ground segmentation (and perhaps even binocular fusion to objects) in a 3-dimensional world that provides rich stereo and ocular contrast information in inputs (Li & Atick, 1994).

Since our gaze is typically controlled by both bottom-up and goal-directed factors, top-down attentional control can be used to limit gaze capture in many cases. For example, a salient, but task-irrelevant, red flower among green leaves can have its influence over gaze be weakened by top-down mechanisms during a task such as searching for blueberries. Favoring the task-relevant blue feature and suppressing the non-target red feature is a form of feature-based top-down attention (Maunsell & Treue, 2006). The feature-based top-down attention can be captured in psychological or phenomenological models as task-dependent weights on various features, with the weighted sum of input strengths from various features giving the overall attentional attraction for any spatial location (Wolfe, 1994). However, one might expect that the higher brain mechanisms would need access to the visual features concerned to exert top-down influences, and so it is an interesting question for future investigations whether this would be less effective for a feature such as ocularity for which awareness is limited. Meanwhile, it has also been proposed that feature dimensional weighting for attentional guidance can occur in early stages of visual processing or pre-attentively (Müller & Krummenacher, 2006), and if so, such weighting could also occur for ocularity. My current findings are only the first forays in design of visual stimuli involving ocularity to explore the link between three important aspects of brain functions: sensory awareness, attentional attraction, and motor action.

## Conclusion

In some situations, a task-irrelevant ocular singleton can capture gaze more strongly than a perceptually more distinctive orientation singleton, which is the target in a speeded visual search task. When the illusory contrast cue associated with the ocular singleton can be exploited to distinguish it from other visual items in a scene, observers

may be aware of its presence. However, in a substantial fraction of trials, observers were unaware of the ocular singleton, even though it captures their gaze, and this degree of awareness was independent of whether the ocular singleton captured gaze. When the illusory contrast cue is overly submerged by inhomogeneous contrasts in the input stimuli, awareness of the ocular singleton was reduced to near chance levels. However, in this case, awareness did become significant when gaze was captured.

## Acknowledgments

This work was supported by the Gatsby Charitable Foundation and British Cognitive Science Foresight Grant BBSRC #GR/E002536/01/. I would like to thank Martin Eimer and his laboratory technician Sue Nicholas for advise on the EEG techniques, the two reviewers for their very helpful comments on this paper, and Peter Dayan for help with English editing.

Commercial relationships: none.

Corresponding author: Li Zhaoping.

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

Address: 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] [Article]
- Belopolsky, A. V., Kramer, A. F., & Theeuwes, J. (2008). The role of awareness in processing of oculomotor capture: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, 12, 2285–2297. [PubMed]
- Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, 33, 1–21. [PubMed] [Article]
- Burkhalter, A., & Van Essen, D. C. (1986). Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey. *Journal of Neuroscience*, 6, 2327–2351. [PubMed] [Article]
- Corbetta, M., & Shulman, G. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3, 201–215. [PubMed]
- Crick, F., & Koch, C. (1995). Are we aware of neural activities in primary visual cortex? *Nature*, 375, 121–123. [PubMed]
- DeAngelis, G. C., Freeman, R. D., & Ohzawa, I. (1994). Length and width tuning of neurons in the cat's primary visual cortex. *Journal of Neurophysiology*, 71, 347–374. [PubMed]
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222. [PubMed]
- Frith, C. D., Blakemore, S. J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 355, 1771–1788. [PubMed]
- Georgeson, M. (1997). Vision and action: You ain't seen nothin' yet ... *Perception*, 26, 1–6. [PubMed]
- Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, 391, 481–484. [PubMed]
- Hasegawa, R. P., Peterson, B. W., & Goldberg, M. E. (2004). Prefrontal neurons coding suppression of specific saccades. *Neuron*, 43, 415–425. [PubMed]
- Hoffman, J. E. (1998). *Visual attention and eye movements*. In H. Pashler (Ed.), *Attention* (pp. 119–154). Philadelphia: Taylor & Francis Press.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195, 215–243. [PubMed]
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2, 194–203. [PubMed]
- 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]
- Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. B. Long & A. D. Baddeley (Eds.), *Attention and Performance IX* (pp. 187–203). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience*, 23, 315–341. [PubMed]
- 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]
- 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://www.journalofvision.org/content/7/7/6>, doi:10.1167/7.7.6. [PubMed] [Article]

- Leach, J. C., & Carpenter, R. H. (2001). Saccadic choice with asynchronous targets: Evidence for independent randomisation. *Vision Research*, *41*, 3437–3445. [[PubMed](#)]
- Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*, 635–654. [[PubMed](#)]
- 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 intracortical interactions in primary visual cortex. *Network: Computation in Neural Systems*, *10*, 187–212. [[PubMed](#)]
- Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, *6*, 9–16. [[PubMed](#)]
- Li, Z., & Atick, J. J. (1994). Efficient stereo coding in the multiscale representation. *Network: Computation in Neural Systems*, *5*, 157–174. [[Article](#)]
- Marg, E. (1951). Development of electro-oculography, standing potential of the eye in registration of eye movement. *A. M. A. Archives of Ophthalmology*, *45*, 169–185. [[PubMed](#)]
- Masson, G. S., Busetini, C., & Miles, F. A. (1997). Vergence eye movements in response to binocular disparity without depth perception. *Nature*, *389*, 283–286. [[PubMed](#)]
- Maunsell, J. H., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neuroscience*, *29*, 317–322. [[PubMed](#)]
- McCormick, P. A. (1997). Orienting attention without awareness. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 168–180. [[PubMed](#)]
- Mulckhuyse, M., Talsma, D., & Theeuwes, J. (2007). Grabbing attention without knowing: Automatic capture of attention by subliminal spatial cues. *Visual Cognition*, *15*, 779–788. [[Article](#)]
- Müller, H. J., & Krummenacher, J. (2006). Locus of dimension weighting: Preattentive or postselective? *Visual Cognition*, *14*, 490–513. [[Article](#)]
- Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, *29*, 1631–1647. [[PubMed](#)]
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42. [[PubMed](#)]
- Schiller, P. H. (1998). The neural control of visually guided eye movements. In J. E. Richards (Ed.), *Cognitive neuroscience of attention, a developmental perspective* (pp. 3–50). London: Lawrence Erlbaum Associates.
- Serences, J. T., & Yantis, S. (2007). Spatially selective representations of voluntary and stimulus-driven attentional priority in human occipital, parietal, and frontal cortex. *Cerebral Cortex*, *17*, 284–293. [[PubMed](#)]
- Shanks, D. R. (2010). Learning: From association to cognition. *Annual Review of Psychology*, *61*, 273–301. [[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](#)]
- Tavassoli, A., & Ringach, D. L. (2010). When your eyes see more than you do. *Current Biology*, *20*, R93–R94. [[PubMed](#)]
- 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](#)]
- Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*, 599–606. [[PubMed](#)]
- Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Psychological Science*, *9*, 379–385. [[Article](#)]
- Thompson, K. G., & Bichot, N. P. (2005). A visual salience map in the primate frontal eye field. *Progress in Brain Research*, *147*, 251–262. [[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](#)]
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238. [[Article](#)]
- Wolfe, J. M., & Franzel, S. L. (1988). Binocularity and visual search. *Perception & Psychophysics*, *44*, 81–93. [[PubMed](#)]
- Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention* (pp. 223–256). Philadelphia: Taylor & Francis Press.
- Zeki, S. M. (1978). Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex. *The Journal of Physiology*, *277*, 273–290. [[PubMed](#)]

- Zhang, X., Zhaoping, L., Zhou, T., & Fang, F. (2012). Neural activities in V1 create a bottom-up saliency map. *Neuron*, *73*, 183–192. [[PubMed](#)]
- Zhaoping, L. (2008a). 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://www.journalofvision.org/content/8/5/1>, doi:10.1167/8.5.1. [[PubMed](#)] [[Article](#)]
- Zhaoping, L. (2008b). After-search—Visual search by gaze shifts after input image vanishes. *Journal of Vision*, *8*(14):26, 1–11, <http://www.journalofvision.org/content/8/14/26>, doi:10.1167/8.14.26. [[PubMed](#)] [[Article](#)]
- Zhaoping, L. (2010a). Gaze capture by task-irrelevant, eye of origin, singletons even without awareness during visual search [[Abstract](#)]. *Journal of Vision*, *10*(7):1318, 1318a, <http://www.journalofvision.org/content/10/7/1318>, doi:10.1167/10.7.1318.
- Zhaoping, L. (2010b). Ocularity as a basic visual feature dimension for bottom-up attentional attraction. *Perception*, *39*, ECVF Abstract Supplement, 4. [[Article](#)]
- Zhaoping, L., & Frith, U. (2011). A clash of bottom-up and top-down processes in visual search: The reversed letter effect revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 997–1006. [[PubMed](#)]
- Zhaoping, L., & Guyader, N. (2007). Interference with bottom-up feature detection by higher-level object recognition. *Current Biology*, *17*, 26–31. [[PubMed](#)]
- 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](#)]