Attention capture by eye of origin singletons even without awareness—A hallmark of a bottom-up saliency map in the primary visual cortex

Li Zhaoping

Department of Computer Science,
University College London, UK

Human observers are typically unaware of the eye of origin of visual inputs. This study shows that an eye of origin or ocular singleton, e.g., an item in the left eye among background items in the right eye, can nevertheless attract attention automatically. Observers searched for a uniquely oriented bar, i.e., an orientation singleton, in a background of horizontal bars. Their reports of the tilt direction of the search target in a brief (200 ms) display were more accurate in a dichoptic congruent (DC) condition, when the target was also an ocular singleton, than in a monocular (M) condition, when all bars were presented to the same single eye, or a dichoptic incongruent (DI) condition, when an ocular singleton was a background bar. The better performance in DC did not depend on the ability of the observers to report the presence of an ocular singleton by making forced choices in the same stimuli (though without the orientation singleton). This suggests that the ocular singleton exogenously cued attention to its location, facilitating the identification of the tilt singleton in the DC condition. When the search display persisted without being masked, observers' reaction times (RTs) for reporting the location of the search target were shorter in the DC, and longer in the DI, than the M condition, regardless of whether the observers were aware that different conditions existed. In an analogous design, similar RT patterns were observed for the task of finding an orientation contrast texture border. These results suggest that in typical trials, attention was more quickly attracted to or initially distracted from the target in the DC or DI condition, respectively. Hence, an ocular singleton, though elusive to awareness, can effectively compete for attention with an orientation singleton (tilted 20 or 50 degrees from background bars in the current study). Similarly, it can also make a difficult visual search easier by diminishing the set size effect. Since monocular neurons with the eye of origin information are abundant in the primary visual cortex (V1) and scarce in other cortical areas, and since visual awareness is believed to be absent or weaker in V1 than in other cortical areas, our results provide a hallmark of the role of V1 in creating a bottom-up saliency map to guide attentional selection.

Introduction

Due to a bottleneck in attentional processes, the visual system can only select a fraction of input for detailed processing, often by shifting gaze in free viewing (Hoffman, 1998). Selection can be according to top-down goals, such as directing gaze to a book in reading, and/or by bottom-up or goal-independent factors, such as when distracted by a sudden movement in the visual periphery (Pashler, 1988, 1998). As bottom-up selection is typically faster (Nakayama & Mackeben, 1989) and often more potent (Jonides, 1981) than top-down selection, understanding it is critical for understanding visual selection as a whole. This paper focuses on bottom-up attraction to attentional selection and uses the term saliency for the strength of this attraction.

It is widely believed that attention is guided by a saliency map which contains a saliency value for each visual location. In the traditional view (Itti & Koch, 2001; Koch & Ullman, 1985; Wolfe, Cave, & Franzel, 1989; Wolfe, 1994), this map is the result of summing activation values from separate feature maps, each of which processes visual inputs associated with a corresponding feature value (like red color, green color, vertical orientation, and rightward motion direction) in one of the basic feature dimensions such as color, orientation, and motion direction (Julesz, 1981; Treisman & Gelade, 1980). This saliency map model, combined with considerations of feature similarities (Duncan & Humphreys, 1989), accounts quite well for substantial behavioral data, such as fast visual searches for feature singletons (e.g., a vertical bar among horizontal ones), slower searches for targets defined by conjunctions of basic features, and pop-out of texture borders defined by, e.g., an orientation contrast (Julesz, 1981; Nothdurft, 1991, 1992, 1993; Treisman & Gelade, 1980; Wolfe, 1998). This traditional view implies that the neurons in the saliency map are not tuned to any basic features, hence suggesting that the map must be in higher cortical areas. This has motivated

searches for this map in areas such as the lateral–intraparietal area (LIP) (Gottlieb, Kusunoki, & Goldberg, 1998) and the frontal eye field (FEF (Schall & Thompson, 1999).

More recently, we have proposed that the primary visual cortex (V1) creates a bottom-up saliency map (Li, 1999a, 1999b, 2002; Zhaoping, 2005), despite the fact that its neurons are tuned to basic input features. In this theory, the most salient location is the spatial receptive field of the most responsive V1 cell to the input scene, regardless of the input selectivity of the neuron, as if the neurons are bidding for visual selection in an auction using their responses as a universal currency (Zhaoping, 2006; Zhaoping & Dayan, 2006). This proposal was partly motivated by the finding that a V1 neuron's response can be significantly suppressed by contextual inputs outside but near its receptive field (e.g., Allman, Miezin, & McGuinness, 1985; Jones, Grieve, Wang, & Sillito, 2001; Kastner, Nothdurft, & Pigarev, 1997, 1999; Knierim & van Essen, 1992; Lamme, 1995; Li & Li, 1994; Nothdurft, Gallant, & van Essen, 1999; Sillito, Grieve, Jones, Cudeiro, & Davis, 1995; Wachtler, Sejnowski, & Albright, 2003): The response to its preferred input feature, in orientation, color, or motion direction etc., is much more suppressed when there are similar rather than very different input features in the nearby context. Specific examples of such iso-feature suppressions include iso-orientation suppression (e.g., Knierim & van Essen, 1992), iso-color suppression (Wachtler et al., 2003), and iso-motion-direction suppression (Jones et al., 2001). The intra-cortical connections linking nearby V1 neurons (Gilbert & Wiesel, 1983; Rockland & Lund, 1983) are believed to mediate the suppression. As an instantiation of the theory, we showed that the responses of a physiologically based model of V1 that incorporates such connections can account for much of the behavioral data on visual searches and segmentation tasks, reflecting bottom-up saliency (Li, 1999a, 1999b, 2000, 2002). In particular, iso-feature suppression makes V1 response highest to feature singletons. For instance, the neuron responding to an orientation singleton in a background of uniformly oriented bars is typically the most responsive neuron since it is the only responding neuron to escape from the iso-orientation suppression experienced by the other neurons responding to the background bars. This most active neuron attracts attention to its receptive field, making the singleton salient.

In addition to explicitly specifying a neural basis for saliency, the V1 saliency hypothesis differs fundamentally from the traditional saliency model by not requiring separate feature maps or any summation of them, and thus not requiring the master map whose neurons are untuned to input features. This difference can be stated in a simplistic way (Zhaoping & May, 2007): under the V1 hypothesis, the saliency value at a location corresponds to the *maximum* of all the responses of the V1 neurons to this location, whereas under the traditional saliency model, it corresponds to the *summation* of the feature map responses to the location. These two rules for computing saliency from responses are termed here the MAX rule (for the V1 saliency hypothesis) and the SUM rule (for the traditional model), respectively, noting that in the former case the responses are from the V1 neurons and in the latter from the units in the feature maps. Consequently, the two different hypotheses for saliency can make qualitatively different predictions about visual selection behavior. Some recent psychophysical tests have confirmed behavioral predictions from the MAX rule of the V1 hypothesis (Koene & Zhaoping, 2007; Zhaoping & May, 2007; Zhaoping & Snowden, 2006; Jingling & Zhaoping, 2008).

The current study aims to distinguish the two hypotheses further by using two well known differences between V1 and other cortical areas. First, relative to any other cortical area (Burkhalter & van Essen 1986; Hubel & Livingstone, 1987; Hubel & Wiesel, 1968; Zeki, 1978), V1 has substantially more monocular cells that contain the information about eye of origin. Recently, it was observed (DeAngelis, Freeman, & Ohzawa, 1994; Webb, Dhruv, Solomon, Tailby, & Lennie, 2005) that the response from a V1 cell to stimulus presented to one eye within its receptive field tends to be suppressed more strongly by contextual input presented to the same rather than the other eye, suggesting that iso-ocular suppression exists in V1 just like iso-feature suppression in orientation, color, or motion direction feature. Hence, the V1 hypothesis makes the following prediction: A visual location should be salient or attract attention automatically when it is at an ocular singleton or ocular contrast, i.e., an item presented uniquely to one eye among uniform background items presented to the other eye or a border between a texture presented to the left eye and another to the right eye. This predicted saliency by ocular discontinuity should hardly be attributed to mechanisms in higher cortical areas. Second, V1 is, perhaps arguably, the cortical area whose neural responses are least correlated with visual awareness (Crick & Koch, 1995; He, Cavanagh, & Intriligator, 1996; He & MacLeod, 2001; for more details, see Discussion section). Indeed, Wolfe and Franzel (1988) showed that human observers cannot report whether an ocular singleton is present in an image, and others (Kolb & Braun, 1995; Morgan, Mason, & Solomon 1997) showed that human observers often lack confidence in locating a texture boundary defined by a change in eye of origin. These observations indicate that information about ocular contrast is elusive to awareness, confirm that ocular information available in V1 is barely available in higher cortical areas associated with awareness, and since saliency and access to awareness are not a prior linked together, do not contradict the V1 prediction above. Meanwhile, saliency by the elusive (to awareness) ocular information would be a hallmark of V1's particular role and indeed would be perhaps the ultimate exogenous cue.

When observers are unaware of an input ocular contrast, probing its effect on saliency requires behavioral tests in which this factor is task irrelevant (so that observers do not have to report it) but can nevertheless be manifest in the accuracy or speed of performance. Here, human observers searched for an orientation singleton, i.e., a uniquely tilted bar among uniformly oriented background bars, or an orientation contrast border between two textures of uniformly tilted bars. In Experiment 1, the task was to report the tilt direction of the orientation singleton from the background bars in an image displayed for only 200 milliseconds (ms). This is very difficult since unless attention was somehow guided to the target, the image was too brief for the target to be properly located and its tilt identified. In Experiments 2–4, the task was to report as soon as possible the location of the orientation singleton or texture border in a stimulus image displayed without a mask until report. In all experiments, some stimuli contained an ocular singleton or ocular contrast border, i.e., a bar presented uniquely to one eye among bars presented to the other eye or a border between a texture of bars presented to the left eye and another of bars to the right eye. The ocular discontinuity (singleton or border) was task irrelevant, and since the tasks were executed quickly or the stimuli presented briefly, the subjects were typically unaware of it unless informed or were unable to identify it even when forced to (Experiment 1B). If, nevertheless, the ocular discontinuity was the most salient location in the display, i.e., more salient than the task-relevant orientation discontinuity (i.e., the target), it should lead attention to the target more quickly when it coincides with the target, or away from it when it is away from the target, manifesting in better/faster or worse/slower performance respectively. These findings are reported next (Zhaoping, 2007a, 2007b).

Methods

Participants

All observers (subjects) were adults between 18 and 45 years old, had normal or corrected-to-normal vision, and, except for LZ (the author) in Experiment 1, were unaware of the research goal of the experiments. All were tested for their ability to see stereo depth as follows. They were shown a regular texture of (22 rows by 30 columns of) vertical bars (each is $0.12° \times 1.1°$ in visual angle) at zero disparity and another (analogous) texture of horizontal bars at 0.6° uncrossed disparity and were asked which bars were in front. They were also asked about the depth of the four corner anchoring points (at zero disparity) relative to the texture bars. This paper reports results only from subjects who answered these questions correctly and can clearly see depth. Even though the experiments in this

study do not involve seeing depth, the depth perception test serves to ensure that the subjects have normal vision in both eyes and that they do not have a lazy eye or any other known or unknown abnormality in monocular vision in any single eye.

Stimuli and procedures

The stimulus was presented on a Clinton Monoray monitor, at a frame rate of 150 Hz, viewed at a distance of 40 cm in a dim room, with the FE-1 shutter goggles from Cambridge Research Systems (www.crsltd.com). The shutter goggles, with 25% open shutter transmission, 100 μs shutter open–close switching time, and a 500:1 ratio for open:close transmission, let left and right eyes view the temporally alternate frames on the screen, so that each eye views 75 out of the 150 frames each second without any sensation of flicker. Each test stimulus display had 660 bars spanning $34° \times 46°$ in visual angle. Each bar was a rectangle of $0.12° \times 1.1°$ in visual angle, sitting on a regular grid of 22 rows by 30 columns. In Experiments 1 and 4, and for subject EC in Experiments 2 and 3, the location of each bar was randomly jittered horizontally and vertically by up to 0.12° in visual angle. The jitter was to prevent the possibility of an accidental wall paper illusion (which indeed was never reported by any subject), and the results do not seem to depend sensitively on this jitter. A bright dot of size $0.12° \times 0.12°$ sat at the center of mass of each group of four neighboring bars. A disk of 0.5° in diameter sat at each of the four outer corners of the rectangular array of texture bars. A fixation stimulus contained a central fixation point of 0.3° in diameter, together with the same four outer disks in the test stimulus. All the disks and dots, the fixation point, and the instruction text such as "press a button for the next trial" were identical in both eyes, serving to anchor vergence on the display screen. These stimuli were 48 cd/m^2 in brightness, and the background was black. Each stimulus bar was presented to one eye only or, in one stimulus condition of Experiments 2 and 3, identically to both eyes. To present a bar of any particular luminance to one eye only, its luminance alternated from zero luminance in one video frame to double luminance in the next. Without the stereo goggles, a bar appeared equally bright on the screen whether it is presented in the monocular or binocular mode. None of the observers who passed our stereo depth test (see Participants section) reported experiencing binocular rivalry or stereo depth with any experimental stimuli.

Experiment 1

Experiment 1 had two parts. In the main part, Experiment 1A, all bars were horizontal except for an orientation singleton (target) tilted 20° clockwise or counterclockwise
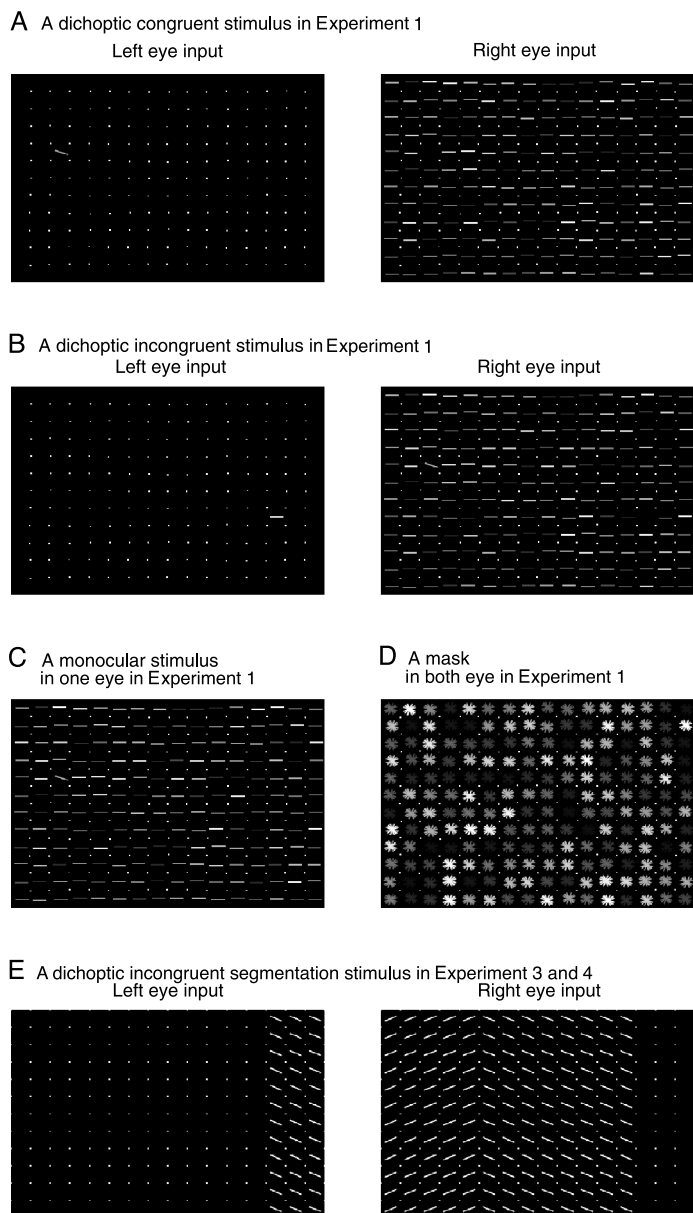
A A dichoptic congruent stimulus in Experiment 1

Left eye input   Right eye input

B A dichoptic incongruent stimulus in Experiment 1

Left eye input   Right eye input

C A monocular stimulus
in one eye in Experiment 1

D A mask
in both eye in Experiment 1

E A dichoptic incongruent segmentation stimulus in Experiment 3 and 4

Left eye input   Right eye input

Figure 1. Illustrative examples A–E of the stimuli. The actual stimuli had 22 rows × 30 columns of bars and had more columns between the orientation discontinuity and the ocular discontinuity in the DI condition (B and E). For the monocular stimulus to one eye in C, the stimulus to the other eye contained the same dots but no bars. In half of the trials of Experiment 1 and all trials in Experiments 2–4, all bars in the test stimulus had the same (uniform) luminance. All test stimulus bars in Experiment 1 were horizontal except the orientation singleton tilted ±20° from horizontal, all those in Experiments 2–4 were tilted ±25° from horizontal.

from horizontal. The target was randomly at one of the 28 (texture) grid locations such that its eccentricity was about 15° and had at least 12° of horizontal eccentricity left or right from the display center. There were three dichoptic

presentation conditions for a test stimulus (see Figure 1): (1) dichoptic congruent (DC), when the target was also an ocular singleton (Figure 1A); (2) dichoptic incongruent (DI), when a background bar on the opposite lateral side of the target from the display center, at one of the same 28 grid locations mentioned above, was an ocular singleton (Figure 1B); and (3) monocular (M), when all bars were presented to the same single eye only (Figure 1C). In each trial, the luminance among the bars was either uniform, for which each bar was 24 cd/m², or non-uniform (as in Figures 1A, 1B, and 1C), for which each bar had a random luminance between 5 and 24 cd/m². The test stimulus was masked binocularly 200 ms after its onset, such that each bar in the test stimulus was replaced in both eyes by a star shaped item, each of which had a random luminance between 2.5 and 24 cd/m² (Figure 1D). The subjects were asked to report by pressing one of the two buttons at their leisure whether the orientation singleton was tilted clockwise or counterclockwise from horizontal, see Figure 2. Experiment 1B, designed to test whether the ocular singleton in Experiment 1A was accessible to awareness, used the same stimuli as Experiment 1A, except that there was no orientation singleton, i.e., all bars were horizontal. An ocular singleton at the same locations as in Experiment 1A was present in half of the trials, and subjects were asked to report at their leisure whether an ocular singleton existed in each trial.

Subjects were asked to view the stimulus without closing either eye. Except for subject LZ (the author), none of the subjects was informed or aware of the task-irrelevant ocular singleton or the existence of different dichoptic (M, DC, and DI) conditions within a data collection session in Experiment 1A. Each subject participated in Experiment 1B after completing Experiment 1A and taking a rest. To ensure that the subjects understood the nature of the ocular singleton in Experiment 1B before performing the task, they were shown an example of the test stimulus containing an ocular singleton for as long a duration as they needed—they were asked to view it by closing one eye at a time and then binocularly to see how the singleton might appear. They were then reminded that

In Experiment 1

Fixation stimulus → Test stimulus → Test stimulus replaced by mask stimulus after 200 ms → Observer reports at leisure the tilt of the orientation singleton (Experiment 1A) or whether an ocular singleton exists (Experiment 1B). → Observer initiates next trial

Time

In Experiment 2–4

Fixation stimulus → Test stimulus → Observer reports the location of the orientation singleton or texture border as soon as possible → Observer initiates next trial
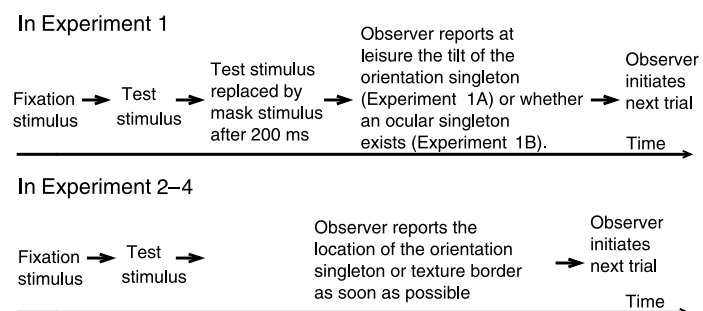
Time

Figure 2. Temporal sequences of events for one trial in Experiment 1 (top) and Experiments 2–4.

they had to open both eyes in each trial during the task proper when data were being taken. Experiment 1A had 2 × 3 task-irrelevant stimulus presentation conditions: (uniform and non-uniform bar luminance) × (M, DC, and DI presentation modes) for tilt singleton identification, and Experiment 1B had 2 task-irrelevant conditions (uniform and non-uniform bar luminance) for ocular singleton detection. There were 90 trials per subject per condition. Each subject performed 24 practice trials (4 trials each condition) in Experiment 1A and 24 practical trials (12 trials each condition) in Experiment 1B before data collection. Data collection for each subject took 40–60 minutes for Experiment 4A, with a short break after every 10–15 minutes, and 15–20 minutes for Experiment 1B, with a short optional break in the middle.

Descriptions in this paragraph apply to all experiments. The stimulus condition in each trial was randomly chosen among all conditions included or stimulus options allowed in the experiment, so that before each trial the subject could not predict beyond chance the presentation condition (e.g., M, DC, or DI), the location of the orientation or ocular singleton or the texture border(s), the tilt of the orientation singleton or the bars in each texture (Experiments 2–4), whether an ocular singleton was present, the eye of origin (randomly left or right in M, DC, and DI conditions) of any bar, whether the bars would have uniform luminance (Experiment 1), and the luminance value at any location when the luminance was non-uniform. Before each trial, an instruction text "press a button for the next trial" was displayed binocularly at the center of the screen. The subject's button press triggered the onset of the fixation stimulus, which was replaced about 1 second later by the test stimulus. Subjects were instructed to fixate on the fixation point before the test stimulus and that they could freely move their eyes after the test stimulus onset. All practice trials were performed immediately before data collection in each session. A beep sounded after each practice trial in all experiments and after every non-practice trial in Experiment 1B (because the task was very difficult), and its pitch provided feedback as to the correctness of the subject's response.

Experiments 2–4

In Experiments 2–4, the task was to find either an orientation singleton or an orientation contrast texture border and to report its location as in the left or right half of the display as soon as possible. The test stimulus stayed on unmasked till the subject's response (see Figure 2). In addition to the M, DC, and DI dichoptic presentation conditions, there was a binocular (B) condition for which all bars were presented identically to both eyes. Different Experiments 2, 3, or 4 differed as to which of the task-irrelevant presentation conditions among B, M, DC, and DI were included in the randomly interleaved trials, and whether observers were informed of them. The task of

finding the orientation singleton or texture border respectively is termed the search or the segmentation task, respectively. Each experimental session had only the search or segmentation task. No subject in Experiments 2–4 participated in Experiment 1.

For the search task, the (monocular) stimuli were the same as that in the Experiment1A, except that (1) each bar had the same luminance of 24 cd/m^2; (2) the target and the background bars were all tilted 25° from horizontal, differing only in the direction of the tilt (clockwise or counterclockwise from horizontal) so that the orientation contrast was 50°; and (3), although of little significance to the purpose of the study, the ocular singleton in the DI condition had an eccentricity randomly between 12° and 15° rather than fixed at 15° in Experiment 1 (due to a programming error). The stimuli for the segmentation task were analogous to those for the search task, such that (1) a vertical texture border defined by an orientation contrast (between bars tilted 25° and −25° from horizontal), as shown for example in Figure 1E, was at 7, 9, or 11 texture columns left or right from the display center; and (2) an ocular contrast vertical border (between bars shown to the left eye and those to the right eye) was present in the DC and DI conditions, coinciding with the orientation texture border in the DC condition, or, in the DI condition, at the opposite lateral side of the orientation texture border, 7, 9, or 11 columns from the display center (see Figure 1E).

To minimize higher-order cognitive effects in the RTs, subjects were instructed to press a button on their left or right hand side, by their left or right hand, respectively, whether the orientation singleton or texture border was in the left or right half of the display. They were also told that they should try not to press the wrong button.

Procedures in experiments for the search task and segmentation task were analogous, so only those for the search task are described here. In Experiment 2, only B, M, and DC conditions were included, and subjects were instructed about the task without being informed about the existence of the task-irrelevant presentation conditions (B, M, and DC) randomly interleaved in the session. Each subject performed 2 practice and 64 non-practice trials in each condition. In Experiment 3, all four conditions B, M, DC, and DI were included, 2 practice and 48 non-practice trials for each condition. All subjects had previously completed Experiment 2 on both the search and segmentation tasks. They were told before data collection that in some trials, they might see a bar that would attract their attention, and that if this bar existed, it would have an equal chance of being the target or a background bar on the opposite lateral side from the target, and that they should try not to be fooled by this distracting bar which may not be the target. Each subject was shown an example of the DI stimuli before data collection and could identify the distracting location as somewhat brighter (without closing any eye) when guided by the author. Experiment 4 was identical to Experiment 2 except that the conditions included were M, DC, and DI, and each subject took

30 practice trials on the M condition and no practice trials for the DC and DI conditions. No subjects in Experiment 4 had participated in Experiments 1–3 before hand (subject EC subsequently participated in Experiments 2 and 3). In Experiments 2, 3, or 4, data collection for each subject and task took one session of 10–15 minutes.

Post-data collection procedures

Immediately after each data collection session of any experiment, the subject was asked for his/her observations and comments and for any strategies used for the task. These questions aimed to find whether observers saw anything unusual such as binocular rivalry or depth, whether they became aware of the task-irrelevant ocular singleton or contrast without being a prior informed about it, and whether they used any strategies such as closing one eye. Note that since the eye of origin (and the tilt) of the orientation singleton or task critical bars was unpredictable in each trial, closing one eye is not an effective strategy. Among all subjects, only one subject, HT in Experiment 4, reported using the strategy in 1–2 trials, before abandoning it after finding it not making the task easier.

In the data analysis, RT outliers were defined as the RTs that were more than three standard deviations different from the mean RT for each subject and stimulus condition in Experiments 2–4. These outliers constitute no more than 4.2% of the trials for each subject and condition. All RT results presented are based on data excluding the practice trials and trials with erroneous responses or outlier RTs. Qualitatively, the same results were obtained when the outliers are retained. All error bars in the figures denote the standard errors of the mean (*SEM*).

Results

Experiment 1: Regardless of its own visibility, an ocular singleton improves identification of an orientation singleton at the same location

Experiment 1 aimed to dissociate the saliency of an ocular singleton (tested in Experiment 1A) from awareness of that singleton. The lack of awareness was apparent in the subjects' inability to detect it in a forced-choice test (Experiment 1B). The detectability of the singleton was made difficult through the use of brief displays of 200 ms (Figure 2) like in previous studies (Kolb & Braun, 1995; Solomon, John, & Morgan 2006; Solomon, Lee, & Sun, 2006) and, in half of the trials, by making the bars have non-uniform luminance values (Figures 1A, 1B, and 1C) like in Wolfe and Franzel's (1988) experiment. As mentioned, subjects had to find a tilt singleton and identify its tilt (in Experiment 1A)—a very difficult task with a

200-ms brief display. The results, shown in Figure 3, indicated that the task performance was better when the tilt singleton coincided spatially with an ocular singleton (DC) than that when the ocular singleton was at a different location (DI) or was absent (M). Reports from all four naive subjects after data collection indicated that they did not notice the different presentation conditions (M, DC, and DI) randomly interleaved within the session, although it was not uncommon for them to comment that some trials seemed much easier than other trials. The author (subject LZ) felt subjectively that in some trials the target shined bright and clear against a dim (and fuzzy or shapeless) background, while other trials were qualitatively more difficult. Furthermore, the difference in performance between the different conditions did not depend on whether or not the ocular singleton could be detected using a forced choice procedure, itself depending on whether the bars have uniform or non-uniform luminance values. These findings suggest that even though the ocular singleton itself may have been invisible to awareness, it acted as an exogenous cue to attract attention to its location within the short duration of the test stimulus. Thus, in condition DC, performance benefited from enhanced attention or sensitivity to the target. Exogenous (and endogenous) cues are known to be able to enhance sensitivity to visual input at cued locations (Nakayama & Mackeben, 1989; Posner, 1980). Similar performance in M and DI conditions is consistent with previous findings that sensitivities in uncued and invalidly cued trials are comparable (Solomon, 2004). Note that stimulus presentation was too brief for a gaze shift before the onset of a mask, so only covert attention could shift about the display. The presentation was also too brief for binocular rivalry had there be non-identical stimuli at the same retinal location in both eyes.

Note that the tilt singleton was tilted 20° from the background bars. This is more than the just noticeable orientation difference (Foster & Ward, 1991) needed for the bar to pop-out perceptually in a typical visual search task when the display stays unmasked. The high error rates for condition M and the lower error rates for condition DC suggest that the orientation singleton alone was insufficiently salient for its location to pop-out in such a brief display, and that this only happened reliably when it was combined with the ocular singleton. By the MAX rule of the V1 saliency hypothesis (see Introduction section), this suggests that the ocular singleton (alone) may be more salient than the orientation singleton (alone) with a 20° contrast.

Experiments 2–4: Ocular discontinuity at or away from target speeds up or slows down RT, respectively

If an ocular singleton attracts attention more than an orientation singleton, it should speed up the visual search
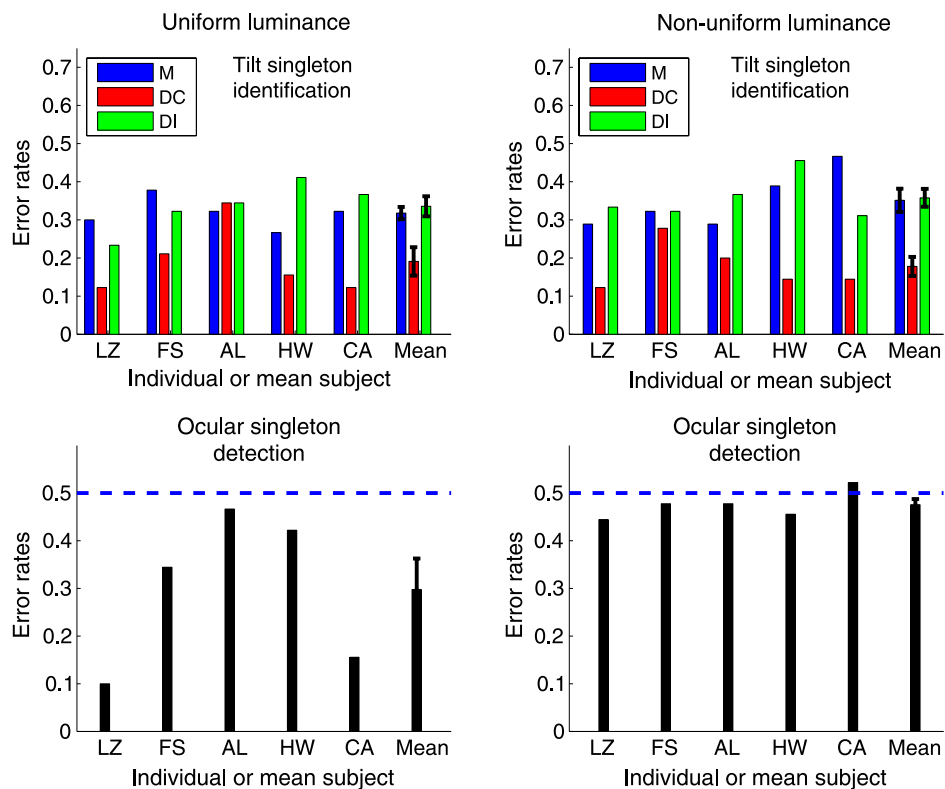
Figure 3. Performance of five subjects LZ, FS, Al, HW, and CA and their overall mean in Experiments 1A (top) and 1B. Blue, red, and green bars display data for M, DC, and DI conditions, respectively (Experiment 1A), and gray bars display data for ocular singleton detection. The proportion of errors in identifying the tilt direction of an orientation singleton in a brief (200 ms) display was significantly ($p < 0.04$) lower when an ocular singleton was at the target location (DC) rather than being elsewhere (DI) or absent (M). When the stimulus bars had uniform or non-uniform luminance values (left or right plots, respectively), the error rates for detecting the ocular singleton itself is significantly ($p = 0.0496$) or not significantly ($p = 0.14$) different from the chance level of 0.5. The error rates in each dichoptic condition, M, DC, or DI, do not ($p > 0.4$) depend significantly on the uniformity of the luminance of the bars. Error rates for M and DI conditions were not significantly different ($p > 0.6$) between luminance conditions. All the p values come from (matched sample) t-test across subjects. Subjects FS, HW, and CA had no previous experience as subjects in visual psychophysics experiments.

for the (orientation singleton) target in the DC condition (when they coincide spatially) and slow down the search in the DI condition (when the former outcompetes the latter, thus distracting attention away from the target during the search). Similarly, a task-irrelevant, but sufficiently salient, color singleton can speed up or slow down a search for orientation singleton according to whether or not it is in the same location as the target (Krummenacher, Müller, & Helller 2001; Pashler, 1988; Zhaoping & May, 2007). Of course, a unique color feature is highly visible to awareness, whereas an ocular feature is not. However, if saliency is dissociated from awareness, the ocular singleton should affect the visual search in just the same way as a color singleton. Experiments 2–4 investigated this by measuring RTs in B, M, DC, and DI conditions (denoted as RT_B, RT_M, RT_{DC}, and RT_{DI}, respectively) in the absence of masking. Based on the argument above, we predict that $RT_{DI} > RT_M > RT_{DC}$. Furthermore, since subjects were instructed to respond as soon as possible, they were likely to mistake the very salient ocular singleton distractor in the DI condition as the target

and thus make an erroneous response. This would manifest itself as differences in error rates across conditions.

Figure 4 shows that indeed $RT_M > RT_{DC}$ ($p < 0.02$) in the visual search task in Experiment 2, even though the subjects were unaware of the presence of the three different task-irrelevant conditions B, M, and DC according to their post-session reports. Since the DI condition was not included in Experiment 2, there was no ocular singleton distractor in any trial. Hence, subjects would never see any background bars as being salient distractors and would not easily suspect or notice the different presentation conditions. Meanwhile, in Experiment 3, which included all B, M, DC, and DI conditions, $RT_{DI} > RT_M$ ($p < 0.04$), and error rates for the DI condition were much higher than for the other conditions, even though subjects were warned of the possible attention grabbing distractor. These results suggest that the effect of the ocular singleton is automatic and not easily turned off by top-down control.

The data are consistent with the following sequence of events in a trial: Attention or gaze is shifted about the
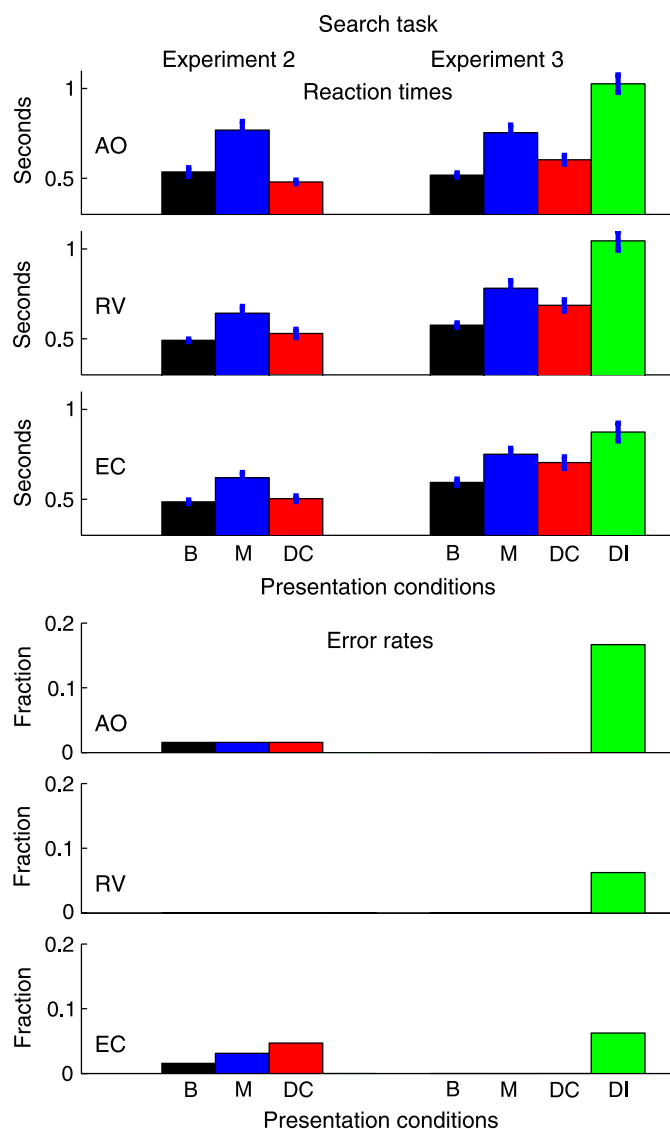
Figure 4. RTs (top) and error rates (the fraction of trials with incorrect button presses) in an orientation singleton search (Experiments 2 and 3) for three naive subjects (AO, RV, and EC). Black, blue, red, or green colored bars indicate binocular (B), monocular (M), dichoptic congruent (DC), or dichoptic incongruent (DI) conditions, respectively. For each subject, RT_{DC} was significantly less than RT_M ($p < 0.02$) in Experiment 2, when the subjects were unaware of the different (B, M, DC) conditions, and RT_{DI} was significantly greater than both RT_M and RT_{DC} ($p < 0.04$) in Experiment 3, when the subjects were informed that some trials might contain a distractor that would grab attention. Averaged across subjects, $RT_{DI} - RT_M = 0.22 \pm 0.05$ second in Experiment 3. In Experiment 3, only subject AO had RT_M significantly greater than RT_{DC} ($p < 0.03$).

scene to find the target. Typically, attention starts at the most salient location, with the observer examining whether the attended item is the target. If it is not, then attention is re-directed to the next most salient location, and so forth, until the target is found. The RT for the trial should reflect the number of attentional shifts required to find the target. Even when the target location is the most salient in the scene, requiring only one attentional shift (from central fixation) to locate it, the RT should decrease with the salience of the target location (Kean & Lambert, 2003). The task-irrelevant ocular singleton competes with the task-relevant orientation singleton for attention. If the ocular singleton is at the same location as the orientation target, then this does not harm performance of the overall task. In fact, if it is more salient than the orientation singleton, the ocular singleton can shorten RT by making attention shift to the target faster. However, when the irrelevant singleton is away from the target and is more salient than the target, attention typically shifts to its location first before being re-directed to the target, and this lengthens RT. Indeed, for trials in which subjects performed correctly, the difference between RT_{DI} and RT_M is about 0.22 second (in Experiment 3 averaged across subjects) or about a fixation duration in typical visual search tasks (Hooge & Erkelens, 1998). This is consistent with the idea that, in most trials in condition DI, the subjects briefly focused their attention on the ocular singleton distractor first and only after realizing that it was not the target did they attend to the orientation singleton.

Even though the ocular contrast is typically not visible to awareness, it is visible to the saliency system that drives attention or saccades. The data again suggest that the ocular singleton is at least as salient as the orientation singleton—note that the orientation singleton here, with a 50° orientation contrast, should be more salient than the orientation singleton in Experiment 1A with a 20° contrast. Furthermore, subjects were not able to exert top-down control to switch off the negative effect of the ocular singleton in DI even when warned ahead of time. This is expected from our knowledge that it is not easy to turn off strong bottom-up saliency factors by top-down control (Jonides, 1981; Yantis & Jonides, 1990), and that the ocular singleton was in any case not strongly present in visual awareness. Since 75% of the trials in Experiment 3 had the target at the location of the most salient bar (for condition B, M, and DC), subjects in a hurry to respond could easily, though mistakenly, report the highly salient ocular singleton location as the target location in condition DI, thus raising error rates. Subjects were apparently more careful in Experiment 3 than Experiment 2, as RT_{DC} is significantly longer in Experiment 3 than that in Experiment 2 across subjects ($p < 0.04$). The uniform luminance of the bars and the longer presentation duration are perhaps the reasons that the subjects were more aware of the ocular singleton in Experiment 3 than Experiment 1.

Since a single input luminance could lead to different apparent brightness in different eyes, the ocular singleton could appear brighter or dimmer than the background bars without any contribution from saliency mechanisms. If this was the reason for a shorter RT_{DC} and a longer RT_{DI}, it should apply only, or at least mainly, when the ocular

singleton was in the eye yielding the higher apparent brightness. Figure 5 shows that the general results above hold more or less regardless of which part of the stimulus was presented to which eye (except that since the number of trials contributing to each data point in the plot is halved by splitting the trials into two different eye of origin categories, some RT differences failed to reach significance). In fact, for each subject, RT_{DC} and RT_{DI} do not significantly vary with eye of origin ($p > 0.05$), even when RT_M does (for subject AO in Experiment 2 only).

Figure 6 shows that the texture segmentation task (performed by the same subjects in the visual search task of Experiments 2 and 3) yields similar results as those from the search task. The high saliency of the ocular texture border should be the basis of the monocular texture border information, which Solomon, John, et al. (2006) and Solomon, Lee, et al. (2006) argued was responsible for ocular contrast-based texture segmentation. It is again apparent that subjects were more cautious in Experiment 3 than Experiment 2. The caution maybe the reason why the RT difference $RT_M - RT_{DC}$ is smaller in Experiment 3 than Experiment 2, though insignificantly ($p = 0.074$, matched sample t-test across subjects and across tasks), and that RTs for the B, M, and DC conditions were also significantly longer ($p < 0.042$) in Experiment 3 for the segmentation task. Despite the caution, the (small) error rates for the B, M, and DC conditions did not change significantly.

When the orientation contrast at the task-relevant location is much less salient by itself (e.g., when the contrast is 20°) than the ocular contrast, such that the RT of the attention shift to it is much longer than that to the ocular contrast, subjects could reduce their RT for the task by adopting the following strategy. If there is a strongly attention capturing location, press the button for this location or the other response button depending on whether there is an orientation contrast at this location, otherwise (in the B and M conditions), search for the orientation contrast to decide which button to press. This strategy would be effective when the orientation contrast at the task-relevant location was sufficiently low, e.g., 20 degrees (data not shown). It would lead to the result (data not shown) $RT_M > RT_{DI}$. In the same way that anti-saccades have longer latencies than saccades, we would still expect $RT_{DC} < RT_{DI}$ (data not shown) for this strategy.

When subjects were not informed about the attention capturing distractor in Experiment 4, the general results obtained in Experiment 3 were still obtained, see Figure 7. None of the subjects in Experiment 4 had participated in any other experiments in this study, and none had been a subject in more than 5 visual psychophysics experiments. For each condition M, DC, and DI, the RT and the error
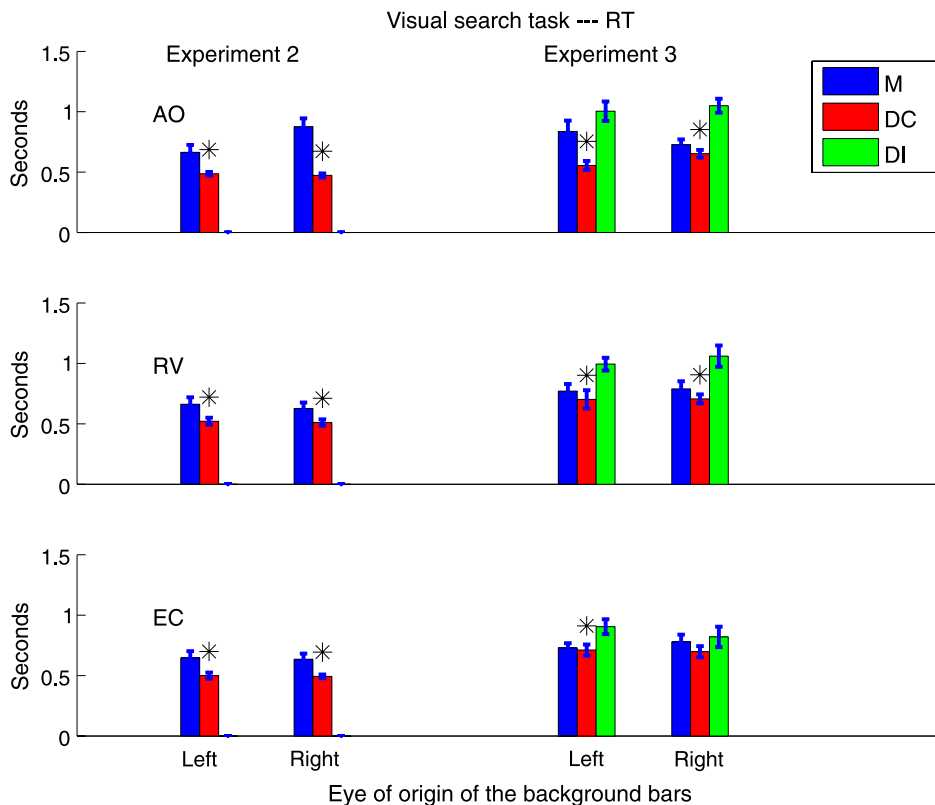

Figure 5. RTs for visual search in Experiments 2 and 3 plotted separately according to the eye of origin of the background bars. "*" above the DC condition data indicate that RT_{DC} is significantly shorter than RT_M in Experiment 2 or RT_{DI} in Experiment 3.
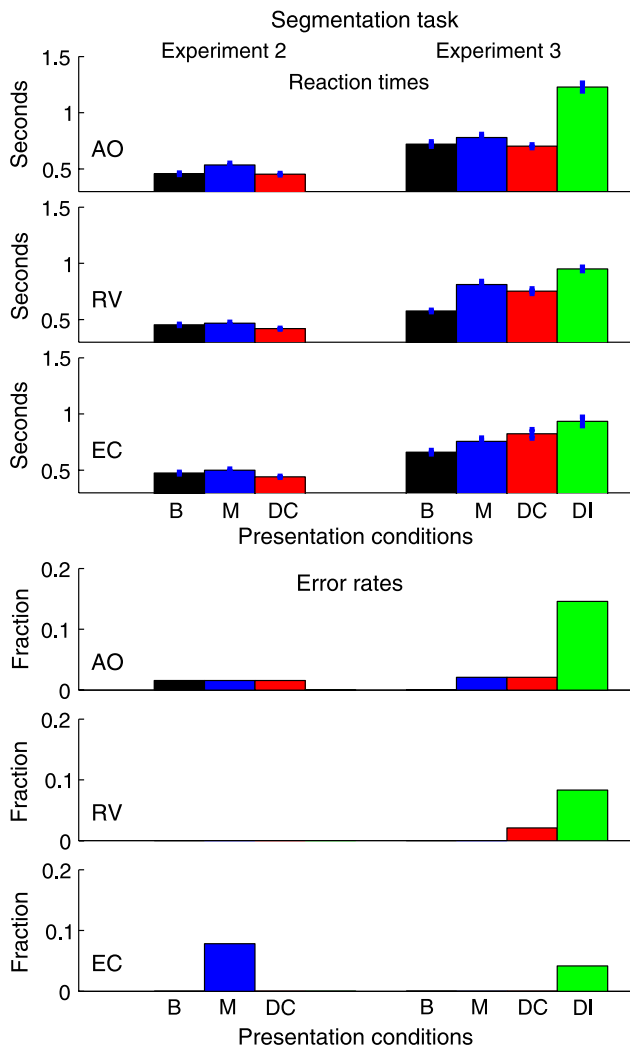
Figure 6. RTs for texture segmentation in Experiments 2 and 3 plotted in the same format as Figure 4 for the search task and involving the same subjects. For each subject, RT_{DC} was significantly less than RT_M ($p < 0.02$) in Experiment 2, and RT_{DI} was significantly greater than both RT_M and RT_{DC} ($p < 0.04$) in Experiment 3 (except for EC for which $RT_{DI} > RT_{DC}$ with $p = 0.052$). In Experiment 3, only subject AO had RT_M significantly greater than RT_{DC} ($p < 0.03$). Across subjects, RTs for B, M, and DC conditions are significantly longer in Experiment 3 than Experiment 2 ($p < 0.042$). Averaged over subjects, $RT_{DI} - RT_M = 0.26 \pm 0.1$ seconds in Experiment 3.

rate across subjects were qualitatively similar to those in Experiment 3 ($p > 0.09$), except for condition M in the segmentation task for which RT was significantly longer in Experiment 3 ($p = 0.04$). Subject DR, who participated in both the search and the segmentation tasks (in different data collection sessions) for this experiment, was the only subject who became aware of the existence of an attention capturing location in some trials, including the fact that it could be at, or away from, the task-relevant location. Subject EC subsequently participated in Experiments 2 and 3. Her RTs across conditions M, DC, and DI for the

segmentation task were significantly longer in Experiment 3 ($p = 0.0153$, matched sample *t*-test), and her error rates for all three conditions were higher in Experiment 4 (barely significant, $p = 0.053$ by matched sample 2-tailed *t*-test). This is consistent with the fact that, in Experiment 3, she was informed of the distracting ocular contrast in the background and was likely more cautious. Averaged across subjects and tasks in Experiment 4, $RT_{DI} - RT_M = 0.18 \pm 0.02$ seconds, suggesting again that in perhaps most trials of condition DI, attention briefly focused on the distracting ocular singleton/contrast location before being re-directed to the task-relevant location. It is unsurprising that six out of the seven subjects were apparently not aware of a distracting location in some trials since shifting attentions about visual space in visual search is common and expected. Once the mind's eye has decided that the attended location has no target, attention departs to continue searching elsewhere and the abandoned location is quickly forgotten (Horowitz & Wolfe, 1998; Zhaoping & Guyader, 2007).

Discussion

Summary of findings

This paper reports that an ocular singleton or discontinuity can automatically attract attention. This is manifest in an improved sensitivity to a briefly displayed target at the location of the ocular singleton, a faster reaction time to a task-relevant location coinciding with the ocular discontinuity, and a slower reaction time when this location was away from that of the ocular discontinuity. These results are consistent with the idea that the ocular discontinuity by being salient and thus automatically capturing attention, guides attention to or away from the task-relevant location. The enhanced sensitivity or shortened RT could occur without the subjects' awareness of (Experiments 1, 2, and 4) or their ability to identify by forced choice (Experiment 1) the presence of this ocular discontinuity. Meanwhile, the ocular discontinuity could distract attention away from the task-relevant location even when the observers were warned not to be distracted by task-irrelevant locations (Experiment 3). This suggests that its effect cannot be easily turned off by top-down control. All these observations suggest that the attraction to attention by the ocular discontinuities is automatic or only under the control of bottom-up factors.

The role of the primary visual cortex in saliency

The findings above strongly implicate the primary visual cortex for such bottom-up saliency, since the eye
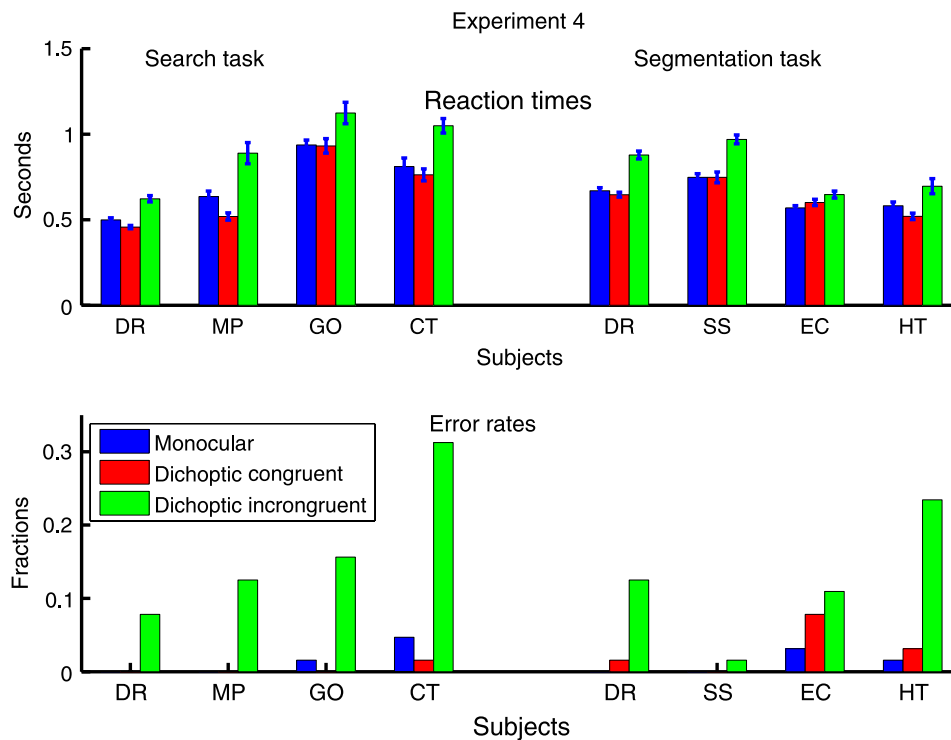
Figure 7. RTs and error rates in Experiment 4 for the visual search (for subjects DR, MP, GO, and CT) and segmentation tasks (for subjects DR, SS, EC, and HT). For each subject, RT_{DI} was significantly greater than RT_M ($p < 0.02$), and, except for subject EC in segmentation task, RT_{DI} was significantly greater than RT_{DC} ($p < 0.02$). In subjects DR and MP for the search task, and HT for the segmentation task, RT_M was significantly greater than RT_{DC} ($p < 0.034$).

of origin information necessary to calculate this saliency requires monocular cells, but is elusive to awareness. First, the primary visual cortex is the only cortical area in which most neurons are monocular cells. In monkeys, more than two thirds (Hubel & Livingstone, 1987; Hubel & Wiesel, 1968) of the V1 cells are classified as monocular (having ocular dominance indices 1, 2, 6, or 7), while monocular cells are only a small percentage of cells in the extra-striate cortices. Zeki (1978) classified 15% of the cells in V2, 17% in V3, 9% in V3A, 6% in V4, and 19% in STS as monocular (according to the current author's estimation from Zeki's plots, and the criteria used to classify a cell's monocularity is not clear from the paper), Burkhalter and Van Essen (1986) classified only 3% of VP cells and 5% of V2 cells as being monocular (defined as cells whose sensitivity to one eye is at least three times of that to the other), and Hubel and Livingstone (1987) classified about (current author's estimation from their plots) 4% of V2 cells as being monocular (having ocular dominance indices 1, 2, 6, or 7). Of course, neurons in the retina and the lateral geniculate nucleus (LGN) are all monocular. However, they are barely tuned to input orientation or direction of motion and hence cannot contribute to saliency by these two features, although they could contribute to saliency by temporal transients, brightness, or color. Our findings could not rule out their role in saliency by ocular contrast.

However, their role would require the mechanism of iso-feature suppression from beyond the classical receptive fields. Physiological observations (Alitto & Usrey, 2008; Allman et al., 1985; Pillow et al., 2007; Solomon, John, et al., 2006; Solomon, Lee, et al., 2006) suggest that such suppression is present only in the magnocellular cells that constitute only a minority of the retinal or LGN output neurons and cannot be wholly responsible for the contextual suppression in V1 (Webb et al., 2005). Furthermore, superior colliculus, which transforms sensory information (such as a saliency map) to gaze shifts, receives inputs from V1 (and other brain sources) and only from the W cells (another small class of cells) from retina, but not from the magnocellular cells or from LGN (Schiller, 1998), suggesting that the role of retina and LGN in saliency is most likely indirect if any.

Second, of all visual cortical areas, the correlation between activities and awareness is considered to be least in V1. Although some functional imaging studies have demonstrated a high correlation between the blood flow signal in V1 and awareness during ambiguous perception under non-ambiguous stimuli (Tong, 2003) or a dissociation between awareness and activity in higher cortical areas (Jiang, Zhou, & He, 2007), there is ample evidence for the contrary conclusion. First, orientation information made inaccessible to awareness by crowding or by being of excessively high spatial frequencies can still be

processed by visual adaptation in V1 (He & MacLeod, 2001; He et al., 1996). Second, V1 responses can follow flicker of color gratings at temporal rates that are too fast to be perceived (Gur & Snodderly, 1997). Third, neuro-imaging studies have found that internally generated forms of visual experience, such as hallucinations (Ffytche et al., 1998), visual auras (Hadjikhani et al., 2001), and color synesthesia evoked by spoken words (Nunn et al., 2002), are associated with blood flow signals in extra-striate cortices but not, or not as immediately, in V1. Fourth, Crick and Koch (1995) reviewed the evidence that V1, in contrast to higher visual areas, is not directly connected to any frontal areas believed to be necessary for reporting awareness. Finally, single cell recordings demonstrated that activities in V1, compared with activities in higher cortical areas, are much less correlated with perception during binocular rivalry under non-changing inputs (Logothetis, Leopold, & Sheinberg, 1996). In sum, it is barely controversial that correspondence between awareness and neural activities is greater in higher cortical areas (Grunewald, Bradley, & Andersen, 2002; Kleinschmidt, Büchel, Zeki, & Frackowiak, 1998). Thus, V1 activities that are selective to the eye of origin of the input should be least likely among the activities in all areas to be accessible by awareness.

The current study adds to the previous studies aimed at testing different hypotheses as to the neural basis of saliency. As mentioned in Introduction section, the traditional idea (Itti & Koch, 2001; Koch & Ullman, 1985; Wolfe et al., 1989) implies that saliency values arise in higher brain areas and result from a SUM rule summing the activations evoked by various input features. Meanwhile, the V1 saliency hypothesis suggests that the saliency of a given location is given by the maximum activation among all V1 neurons favoring this location and responding to any feature(s). A contrasting prediction from this MAX rule is that task-irrelevant features, even when they would lead to a spatially uniform contribution to the master activation map when summed (by the SUM rule of the traditional model), could interfere with visual search or segmentation tasks depending only on task-relevant features. For example, Zhaoping and May (2007) confirmed that segmenting one regular texture consisting of left tilted bars from another consisting of right tilted bars can be made drastically more difficult by the presence of a superposing and task-irrelevant checkerboard texture pattern made of horizontal and vertical bars, with one task-irrelevant bar on each of the original texture bars. Such a prediction would not hold according to the SUM rule. Further, Koene and Zhaoping (2007) provided evidence favoring V1 over V2 as the saliency substrate for the very salient basic feature singleton pop-out. They found that the RT to find a double feature singleton bar that was unique in both color and motion direction could be predicted from a RACE model (according to the MAX rule) between an RT for a color (only) singleton and another for a motion direction (only) singleton. In other words, they found no redundancy gain

in saliency for a redundant feature singleton (unique in both color and motion direction) from the corresponding single feature singleton (unique in color or motion direction only). This finding suggests that, according to the MAX rule, cortical areas responsible for the neural responses (from which to select the maximum response for saliency) should have few neurons tuned conjunctively to both color and motion direction or should receive no inputs from such neurons. This implicates V1 since it has few such neurons (Horwitz & Albright, 2005), whereas V2 and V3 have more (Gegenfurtner Kiper, & Fenstemaker, 1996; Gegenfurtner, Kiper, & Levitt, 1997; Tamura, Sato, Katsuyama, Hata, & Tsumoto, 1996; Shipp, private communication, 2007). Findings from these studies provide converging evidences for the role of V1 in saliency. However, whether and how much higher cortical areas contribute additionally to computing bottom-up saliency is an empirical question to be answered in future studies.

According to the V1 saliency hypothesis or the MAX rule, the enhanced target saliency in the DC condition does not arise because the saliency by ocular contrast and saliency by orientation contrast at the same location are summed up (as in the traditional saliency models Itti & Koch, 2001; Koch & Ullman, 1985; Wolfe et al., 1989). Rather, it arises because the maximum V1 response (among responses from cells tuned to orientation, eye of origin, or both) to the orientation-ocular double feature singleton is higher than the maximum V1 response (from cells tuned to orientation) to the orientation single feature singleton (assuming that cells responding to non-singleton features are much less active). Since RT is reduced in the DC condition, the enhanced maximum V1 response to the double feature singleton should be attributed to a cell tuned to eye of origin. As an ocular singleton also lengthens RT by about one fixation duration in visual search, our findings suggest that saliency by ocular contrast feature alone should be stronger than saliency by the orientation contrasts used in this study.

The iso-ocular suppression

As mentioned in the Introduction section, iso-ocular suppression (i.e., a stronger suppression to a V1 neuron's response to a monocular input within its receptive field from contextual inputs to the same rather than the different eye) should be the mechanism for saliency by ocular contrast. Iso-ocular suppression in V1, studied by DeAngelis et al. (1994) and Webb et al. (2005) and consistent to findings by Macknik and Martinez-Conde (2004), is relatively less known compared to iso-orientation suppression. It is likely that nearby V1 neurons tuned to the same eye or in the same ocular dominance column are more likely linked by intra-cortical connections mediating mutual suppression, just like neurons tuned to the similar orientation are more likely linked by these connections (Bosking, Zhang, Schofield, & Fitzpatrick,

1997). The iso-ocular suppression is consistent with the observation that ocular contrast texture border is more easily or confidently detected in a denser texture (Solomon, John, et al., 2006; Solomon, Lee, et al., 2006)—the intra-cortical connections in V1 that mediate iso-feature suppression only extend up to a few millimeters (Rockland & Lund, 1983), so only input items close enough to each other, i.e., in a denser texture array, can be affected by it. Indeed, visual search for unique basic features, such as a color or orientation singleton, are also easier in denser arrays of (homogenous) background items. Iso-ocular suppression may also be the basis of the observation that the perceived contrast of a (dynamic) texture patch is reduced by a background (dynamic) texture in the same but not in the different eyes (Chubb, Sperling, & Solomon, 1989).

Dissociation between awareness and attraction to attention

Computationally, the elusiveness of eye of origin information to awareness can be understood since the brain's internal model of the contents of the visual world that account for the retinal input typically excludes it (Dayan, 1998), unless it is associated directly with depth. However, this elusiveness may explain why the high salience of ocular contrast was not hitherto noticed. In previous studies (Kolb & Braun, 1995; Morgan et al., 1997, Solomon, John, et al., 2006; Solomon, Lee, et al., 2006; Wolfe & Franzel, 1988), the main concern was whether ocular contrast alone, without other cues, was reportable, i.e., visible to awareness, rather than whether or not it could attract attention. Hence, saliency of an ocular singleton that was barely visible to awareness would be difficult to assess in tasks requiring its presence to be reported. Indeed, failure consciously to recognize a target initially selected only by bottom-up saliency has been observed explicitly in a recent study in visual search using eye tracking (Zhaoping & Guyader, 2007). Bottom-up saliency arising from an orientation singleton, a low level feature, led gaze to a target item, but then an attentive process judged the selected item not to be different from the distractors in its higher level object shape, and so the gaze abandoned the target to continue searching elsewhere. In searching for an ocular singleton without other cues, bottom-up saliency by ocular contrast should lead attention to the target, but higher level attention would see the ocular singleton as not being distinct from distractors in its brightness or shape. The current study unveils the automatic attraction to attention by making the ocular signal task-irrelevant and observing its facilitation or interference in another task that requires attention to the task-relevant location.

Empirically, if a feature singleton pops-out or has little set size effect (i.e., RT for search does not increase with the number of background items) in visual search, the feature is defined as a basic feature (Treisman & Gelade,

1980). Orientation is one such basic feature. The current study suggests that the ocular singleton is at least as salient as and most likely more salient than the orientation singleton that has an orientation contrast of 20 degrees (Experiment 1A) or even 50 degrees (Experiments 2–4), since it could apparently robustly compete with the orientation singleton for attention in the DI condition and significantly reduce the RT and improve performance for orienting to the spatially coinciding orientation single-ton in the DC condition. Hence, the ocular feature should be a basic feature if it could be searched for and reported, i.e., when it could be accessed by awareness. One can predict that searching for a non-basic feature singleton can be made easier if the target is also an ocular singleton. Indeed, the set size effect in searching for a side-way letter T among letter L's is removed by making the T an ocular singleton, see Figure 8.

Although ocular contrast could be invisible to aware-ness, according to this study, it is highly visible to the saliency or the oculomotor system. It is well known that post-selectional processing and awareness can be disso-ciated (Koch & Tsuchiya, 2007). For instance, one can
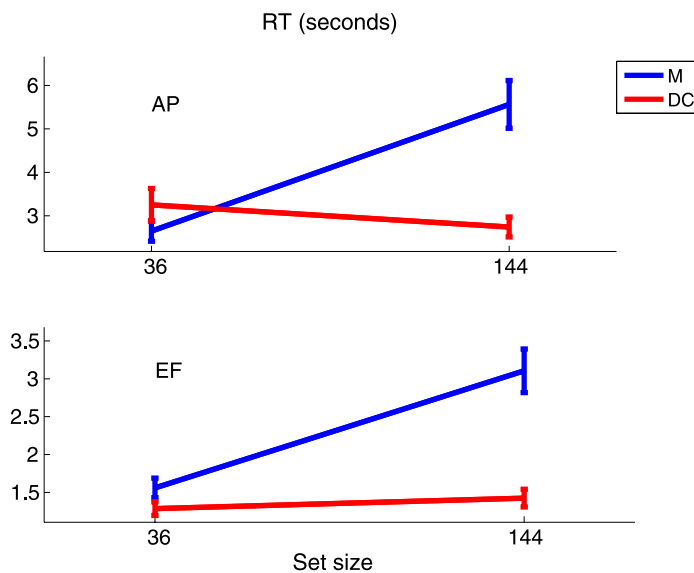

Figure 8. An ocular singleton removes the set size effect in searching for a side way letter T among letter L's in two subjects AP and EF. The task was to report quickly whether the T, present in each trial, was pointing to the left or right. The stimuli were designed like the test stimuli in Experiments 2–4 except that the set size or the total number of items in the display was varied (by varying the density of items), and all items were displayed within a binocular square frame of size about $28° \times 28°$. Trials of different set sizes and presentation conditions M and DC were randomly interleaved. Subjects were unaware of the different dichoptic conditions and were instructed to search by looking about the display randomly rather than strategically, (e.g., to avoid search-ing line by line in apparently difficult searches) to inhibit top-down control from overriding bottom-up saliency.

pay attention to a letter within a letter stream in the visual periphery without being able to recognize it; however, one is perfectly aware of the fact that attention is being directed to that location or that this location is selected by attention. Dissociation between attentional selection (with perceptually measurable effects) and awareness as seen in our study, when subjects were unaware that attention was attracted to a location, is a more recently discovered phenomenon. Jiang, Costello, Fang, Huang, and He (2006) showed another example of dissociation between attentional selection and awareness: Inter-ocularly suppressed, and thus invisible, erotic images attracted or repelled attention and thus affected discrimination performance based on visual stimuli presented very soon afterwards. However, this selectional effect was only significant at a group level, whereas in our study, it was much stronger, being significant in individual subjects. This dissociation between selection and awareness might seem surprising. However, it is somewhat analogous to the dissociation that has been observed between object recognition and visually guided grasping, with patients with particular brain lesions being able to orient their hands correctly with respect to an object, without recognizing the object or its orientation, i.e., action without perception (Georgeson, 1997; Milner & Goodale, 1995). As the functional role of saliency is to guide attention to the salient locations, the saliency computation belongs to computations on visual inputs associated with "where" and not "what" (Sagi & Julesz, 1985; Zhaoping & May, 2007). After all, if saliency had the "what" information there would be no need to direct attention to the location to rediscover "what". The saliency computed in V1 can directly affect behavior through V1's monosynaptic connection to the superior colliculus that controls eye movements (Fecteau, Bell, & Munoz, 2004; Tehovnik, Slocum, & Schiller, 2003). Ocular contrast tends to be present at the boundaries between surfaces of different depths. High saliency by ocular contrast can thus help to segment a foreground from a background by directing attention to the depth boundary between them. Once this segmentation is achieved (and perhaps also after the depth order between surfaces is obtained), the brain may consider it unnecessary to explicitly retain the information about eye of origin or ocular contrast in higher brain areas or subsequent stages of processing. Experiment 1B suggests that the saliency effect of the ocular singleton is brief—even though it can attract attention in a briefly displayed stimulus, this attraction apparently did not last long enough to cue the observers for reporting the presence of this attraction (i.e., for it to enter awareness) in an unspeeded forced choice task. This is consistent with previous findings on the transient nature of exogenous attention (Nakayama & Mackeben, 1989; van Zoest & Donk, 2006). The brevity of the saliency effect is consistent with the functional role of salience to guide attention; after all, this guidance is no longer needed once attention has followed it to the salient location, or has

decided to ignore it in a typical visual environment. Note that Experiment 1B revealed that observers did not even have the blind sight (defined as competent stimulus identification coupled with unconfident reports) with sufficient non-uniformity in luminance, even though they could have it in other stimulus conditions (Kolb & Braun, 1995). Additionally, the saliency signals could also be used in higher areas such as the extra-striate cortex (Beck & Kastner, 2005; Reynolds & Desimone, 2003), LIP (Gottlieb et al., 1998), and FEF, to combine with top-down controls for visual processing.

Conclusion

The current study showed that inputs of unique eye of origin or ocular discontinuities can automatically attract attention even when the subjects are not aware of it, are actively avoiding it, or are unable to identify the ocular contrast in forced choice tasks. Since, among the visual cortical areas, V1 contains the most monocular cells for ocular information and is the least associated with visual awareness, these findings provide a hallmark of the primary visual cortex in its role of generating a bottom-up saliency map to guide attention.

The current findings, together with previous studies (Koene & Zhaoping, 2007; Zhaoping & May, 2007; Zhaoping & Snowden, 2006; Jingling & Zhaoping, 2008), provide converging evidence that bottom-up saliency could be computed at a lower level visual area than traditionally thought, and that there is no need for a master saliency map for the bottom-up saliency (even if another cortical area may be needed for the subsequent integration with top-down attentional factors). They also support the proposal that V1's neural responses can be used as universal currency to bid for attentional selection, despite the feature tuning of the V1 neurons (Zhaoping, 2006).

Acknowledgments

Commercial relationships: none.
Corresponding author: Li Zhaoping.
Email: z.li@ucl.ac.uk.
Address: Department of Computer Science, University College London, Gower St., London WC1E 6BT, UK.

Zhaoping 15

References

Alitto, H. J., & Usrey, W. M. (2008). Origin and dynamics of extraclassical suppression in the lateral geniculate nucleus of the macaque monkey. *Neuron, 57,* 135–146. [PubMed]

Allman, J., Miezin, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparison in visual neurons. *Annual Review of Neuroscience, 8,* 407–430. [PubMed]

Beck, D. M., & Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature Neuroscience, 8,* 1110–1116. [PubMed] [Article]

Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *Journal of Neuroscience, 17,* 2112–2127. [PubMed] [Article]

Burkhalter, A., & Van Essen, D. C. (1986). Processing of color, form and disparity information in visual areas VP and V2 of ventral exrastriate cortex in the macaque monkey. *Journal of Neuroscience, 6,* 2327–2351. [PubMed] [Article]

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences of the United States of America, 86,* 9631–9635. [PubMed] [Article]

Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex. *Nature, 375,* 121–123. [PubMed]

Dayan, P. (1998). A hierarchical model of binocular rivalry. *Neural Computation, 10,* 1119–1135. [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]

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458. [PubMed]

Fecteau, J. H., Bell, A. H., & Munoz, D. P. (2004). Neural correlates of automatic and goal-driven biases in orienting spatial attention. *Journal of Neurophysiology, 92,* 1728–1737. [PubMed] [Article]

Ffytche, D. H., Howard, R. J., Brammer, M. J., David, A., Woodruff, P., & Williams, S. (1998). The anatomy of conscious vision: An fMRI study of visual hallucinations. *Nature Neuroscience, 1,* 738–742. [PubMed] [Article]

Foster, D. H., & Ward, P. A. (1991). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proceedings of the Royal Society B: Biological Sciences, 243,* 75–81. [PubMed]

Gegenfurtner, K. R., Kiper, D. C., & Fenstemaker, S. B. (1996). Processing of color, form, and motion in macaque area V2. *Visual Neuroscience, 13,* 161–172. [PubMed]

Gegenfurtner, K. R., Kiper, D. C., & Levitt, J. B. (1997). Functional properties of neurons in macaque area V3. *Journal of Neurophysiology, 77,* 1906–1923. [PubMed] [Article]

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]

Gilbert, C. D., & Wiesel, T. N. (1983). Clustered intrinsic connections in cat visual cortex. *Journal of Neuroscience, 3,* 1116–1133. [PubMed] [Article]

Grunewald, A., Bradley, D. C., & Andersen, R. A. (2002). Neural correlates of structure-from-motion perception in macaque V1 and MT. *Journal of Neuroscience, 22,* 6195–6207. [PubMed] [Article]

Gur, M., & Snodderly, D. M. (1997). A dissociation between brain activity and perception: chromatically opponent cortical neurons signal chromatic flicker that is not perceived. *Vision Research, 37,* 377–382. [PubMed]

Hadjikhani, N., Sanchez Del Rio, M., Wu, O., Schwartz, D., Bakker, D., Fischl, B., et al. (2001). Mechanisms of migraine aura revealed by functional MRI in human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 98,* 4687–4692. [PubMed] [Article]

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383,* 334–337. [PubMed]

He, S., & MacLeod, D. I. (2001). Orientation-selective adaptation and tilt after-effect from invisible patterns. *Nature, 411,* 473–476. [PubMed]

Hoffman, J. E. (1998). In H. Pashler (Ed.), *Visual attention and eye movements* (pp. 119–154). Hove: Psychology Press.

Hooge, I. T., & Erkelens, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research, 38,* 1295–1302. [PubMed]

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394,* 575–577. [PubMed]

Horwitz, G. D., & Albright, T. D. (2005). Paucity of chromatic linear motion detectors in macaque V1. *Journal*

of Vision, 5(6):4, 525–533, http://journalofvision.org/5/6/4/, doi:10.1167/5.6.4. [PubMed] [Article]

Hubel, D. H., & Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area 18. *Journal of Neuroscience, 7,* 3378–415. [PubMed] [Article]

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology, 195,* 215–243. [PubMed] [Article]

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews, Neuroscience, 2,* 194–203. [PubMed]

Jones, H. E., Grieve, K. L., Wang, W., & Sillito, A. M. (2001). Surround suppression in primate V1. *Journal of Neurophysiology, 86,* 2011–2028. [PubMed] [Article]

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender-and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 17048–17052. [PubMed] [Article]

Jiang, Y., Zhou, K., & He, S. (2007). Human visual cortex responds to invisible chromatic flicker. *Nature Neuroscience, 10,* 657–662. [PubMed]

Jingling, L., & Zhaoping, L. (2008). Change detection is easier at texture border bars when they are parallel to the border: Evidence for V1 mechanisms of bottom-up salience. *Perception, 37,* 197–206.

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 Inc.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature, 290,* 91–97. [PubMed]

Kastner, S., Nothdurft, H. C., & Pigarev, I. N. (1997). Neuronal correlates of pop-out in cat striate cortex. *Vision Research, 37,* 371–376. [PubMed]

Kastner, S., Nothdurft, H. C., & Pigarev, I. N. (1999). Neuronal responses to orientation and motion contrast in cat striate cortex. *Visual Neuroscience, 16,* 587–600. [PubMed]

Kean, M., & Lambert, A. (2003). The influence of a salience distinction between bilateral cues on the latency of target detection saccades. *British Journal of Psychology, 94,* 373–388. [PubMed]

Kleinschmidt, A., Büchel, C., Zeki, S., & Frackowiak, R. S. (1998). Human brain activity during spontaneously reversing perception of ambiguous figures. *Proceedings of the Royal Society B: Biological Sciences, 265,* 2427–2433. [PubMed] [Article]

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences, 11,* 16–22. [PubMed]

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4,* 219–227. [PubMed]

Koene, A. R., & Zhaoping, L. (2007). Feature-specific interactions in salience from combined feature contrasts: Evidence for a bottom-up saliency map in V1. *Journal of Vision, 7*(7):6, 1–14, http://journalofvision.org/7/7/6/, doi:10.1167/7.7.6. [PubMed] [Article]

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]

Kolb, F. C., & Braun, J. (1995). Blindsight in normal observers. *Nature, 377,* 336–338. [PubMed]

Krummenacher, J., Müller, H. J., & Heller, D. (2001). Visual search for dimensionally redundant pop-out targets: Evidence for parallel-coactive processing of dimensions. *Perception & Psychophysics, 63,* 901–917. [PubMed] [Article]

Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience, 15,* 1605–1615. [PubMed] [Article]

Li, Z. (1999a). Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proceedings of the National Academy of Sciences of the United States of America, 96,* 10530–10535. [PubMed] [Article]

Li, Z. (1999b). Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network, 10,* 187–212. [PubMed]

Li, Z. (2000). Pre-attentive segmentation in the primary visual cortex. *Spatial Vision, 13,* 25–50. [PubMed]

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences, 6,* 9–16. [PubMed]

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]

Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivaling during binocular rivalry? *Nature, 380,* 621–624. [PubMed]

Macknik, S. L., & Martinez-Conde, S. (2004). Dichoptic visual masking reveals that early binocular neurons exhibit weak interocular suppression: Implications for binocular vision and visual awareness. *Journal of Cognitive Neuroscience, 16,* 1049–1059. [PubMed]

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* Oxford: Oxford University Press.

Morgan, M. J., Mason, A. J., & Solomon, J. A. (1997). Blindsight in normal subjects? *Nature, 385,* 401–402. [PubMed]

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29,* 1631–1647. [PubMed]

Nothdurft, H. C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research, 31,* 1073–1078. [PubMed]

Nothdurft, H. C. (1992). Feature analysis and the role of similarity in pre-attentive vision. *Perception & Psychophysics, 52,* 335–375. [PubMed]

Nothdurft, H. C. (1993). The role of features in pre-attentive vision: Comparison of orientation, motion and color cues. *Vision Research, 33,* 1937–1958. [PubMed]

Nothdurft, H. C., Gallant, J. L., & van Essen, D. C. (1999). Response modulation by texture surround in primate area V1: Correlates of 'popout' under anesthesia. *Visual Neuroscience, 16,* 15–34. [PubMed]

Nunn, J. A., Gregory, L. J., Brammer M., Williams, S. C., Parslow, D. M., Morgan, M. J., et al. (2002). Functional magnetic resonance imaging of synesthesia: Activation of V4/V8 by spoken words. *Nature Neuroscience, 5,* 371–375. [PubMed] [Article]

Pashler, H. (1988). Cross-dimensional interaction and texture segregation. *Perception & Psychophysics, 43,* 307–318. [PubMed]

Pashler, H. (Ed.) (1998). *Attention.* Hove: Psychology Press.

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., et al. (2007). *Correlations and coding with multi-neuronal spike trains in primate retina.* Abstract for the Annual Meeting for Society for Neuroscience, 768.9.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25. [PubMed]

Reynolds, J. H., & Desimone, R. (2003). Interacting roles of attention and visual salience in V4. *Neuron, 37,* 853–863. [PubMed] [Article]

Rockland, K. S., & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *Journal of Comparative Neurology, 216,* 303–318. [PudMed]

Sagi, D., & Julesz, B. (1985). "Where" and "what" in vision. *Science, 228,* 1217–1219. [PubMed]

Schall, J. D., & Thompson, K. G. (1999). Neural selection and control of visually guided eye movements. *Annual Review of Neuroscience, 22,* 241–259. [PubMed]

Schiller, P. (1998). The neural control of visually guided eye movements. In J. E. Richards (Ed.), *Cognitive neuroscience of attention* (pp. 3–50). Mahwah, NJ: Lawrences Erlbaum Associates, Publishers.

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]

Solomon, J. A. (2004). The effect of spatial cues on visual sensitivity. *Vision Research, 44,* 1209–1216. [PubMed]

Solomon, J. A., John, A., & Morgan, M. J. (2006). Monocular texture segmentation and proto-rivalry. *Vision Research, 46,* 1488–1492. [PubMed]

Solomon, S. G., Lee, B. B., & Sun, H. (2006). Suppressive surrounds and contrast gain in magnocellular-pathway retinal ganglion cells of macaque. *Journal of Neuroscience, 26,* 8715–8726. [PubMed] [Article]

Tamura, H., Sato, H., Katsuyama, N., Hata, Y., & Tsumoto, T. (1996). Less segregated processing of visual information in V2 than in V1 of the monkey visual cortex. *European Journal of Neuroscience, 8,* 300–309. [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]

Tong, F. (2003). Primary visual cortex and visual awareness. *Nature Reviews, Neuroscience, 4,* 219–229. [PubMed]

Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97–136. [PubMed]

Van Zoest, W., & Donk, M. (2006). Saccadic target selection as a function of time. *Spatial Vision, 19,* 61–76. [PubMed]

Wachtler, T., Sejnowski, T. J., & Albright, T. D. (2003). Representation of color stimuli in awake macaque primary visual cortex. *Neuron, 37,* 681–691. [PubMed] [Article]

Webb, B. S., Dhruv, N. T., Solomon, S. G., Tailby, C., & Lennie, P. (2005). Early and late mechanisms of surround suppression in striate cortex of Macaque. *Journal of Neuroscience, 25,* 11666–11675. [PubMed] [Article]

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1,* 202–238.

Wolfe, J. M. (1998). In H. Pashler (Ed.), *Visual search, a review* (pp. 13–74). Hove, East Sussex, UK: Psychology Press.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model of visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419–433. [PubMed]

Wolfe, J. M., & Franzel, S. L. (1988). Binocularity and visual search. *Perception & Psychophysics, 44,* 81–93. [PubMed]

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 121–134. [PubMed]

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] [Article]

Zhaoping, L. (2005). The primary visual cortex creates a bottom up saliency map. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (chap. 93, pp. 570–575). San Diego, CA: Elsevier.

Zhaoping, L. (2006). Theoretical understanding of the early visual processes by data compression and data selection. *Network, 17,* 301–334. [PubMed]

Zhaoping, L., & Dayan, P. (2006). Pre-attentive visual selection. *Neural Networks, 19,* 1437–1439. [PubMed]

Zhaoping, L. (2007a). Popout by unique eye of origin: A fingerprint of the role of primary visual cortex in bottom-up saliency. Program No. 717.8 at the Annual Meeting of Society for Neuroscience, San Diego, CA.

Zhaoping, L. (2007b). Unique eye of origin attracts attention automatically even when it cannot be detected by forced-choice—evidence for the role of the primary visual cortex in bottom-up visual saliency. Presented at the twelveth AVA (Applied Vision Association) Christmas meeting, December 17, 2007, at Aston University, Birmingham, UK.

Zhaoping, L., & Guyader, N. (2007). Interference with bottom-up feature detection by higher-level object recognition. *Current Biology, 17,* 26–31. [PubMed] [Article]

Zhaoping, L., & May, K. A. (2007). Psychophysical tests of the hypothesis of a bottom-up saliency map in primary visual cortex. *PLoS Computational Biology, 3,* e62. [PubMed] [Article]

Zhaoping, L., & Snowden, R. J. (2006). A theory of a saliency map in primary visual cortex (V1) tested by psychophysics of color-orientation interference in texture segmentation. *Visual cognition, 14,* 911–933.