

# Learning where to direct gaze during change detection

Department of Brain and Cognitive Sciences and Center for  
Visual Science, University of Rochester,  
Rochester, NY, USA, &

Department of Psychology, University of California  
Santa Barbara, Santa Barbara, CA, USA



**Jason A. Droll**

Department of Brain and Cognitive Sciences and Center for  
Visual Science, University of Rochester, USA, &  
Johns Hopkins University, Baltimore, MD, USA



**Krista Gigone**

Department of Brain and Cognitive Sciences and Center for  
Visual Science, University of Rochester, USA, &  
University of Texas at Austin, Austin, TX, USA



**Mary M. Hayhoe**

Where do observers direct their attention in complex scenes? Previous work on the cognitive control of fixation patterns in natural environments suggests that subjects must learn where to direct attention and gaze. We examined this question in the context of a change blindness paradigm, where some objects were more likely to undergo a change in orientation than others. The experiments revealed that observers are capable of learning the frequency with which objects undergo a change, and that this learning is manifested in the distribution of gaze among objects in the scene, as well as in the reaction time for detecting visual changes, and the frequency of localizing changing objects. However, observers were much less sensitive to the conditional probability of a second feature, border color, predicting a change in orientation. We conclude that striking demonstrations of change blindness may reflect not only the constraints of attention and working memory, but also what observers have learnt about what information to attend and select for storage during the task of change detection. Such exploitation of the frequency of change suggests that gaze allocation is sensitive to the probabilistic structure of the environment.

Keywords: eye movements, gaze control, change detection, change blindness, learning, scenes

Citation: Droll, J. A., Gigone, K., & Hayhoe, M. M. (2007). Learning where to direct gaze during change detection. *Journal of Vision*, 7(14):6, 1–12, <http://journalofvision.org/7/14/6/>, doi:10.1167/7.14.6.

## Introduction

The purpose of this paper is to explore the hypothesis that the control of gaze and resultant sampling of visual information is a consequence of learning the probabilistic structure of the environment. Previous accounts of gaze control often emphasize the low-level attributes in the scene that might capture attention by virtue of their image properties, such as color and chromaticity (Itti & Koch, 2000, 2001) or abrupt onsets (Yantis & Jonides, 1984, 1990). However, these influences are strongly modulated by top-down influences such as the ongoing task (Gibson & Jiang, 1998). Fixations are tightly linked in space and time to the immediate task needs, and few task-irrelevant fixations are observed (Hayhoe, 2000; Hayhoe & Ballard, 2005; Land, Mennie, & Rusted, 1999). The observers' task set that guides such selective gaze control includes not only behavioral goals, but also what information is necessary to acquire in order to achieve those goals. Knowledge of what information is relevant may include

spatial location, or object properties. An implication of the cognitive control of gaze is that observers must learn where to look to get the information they need. This means they must learn the properties of the world and how those properties change (Hayhoe, Droll, & Mennie, 2007; Land & Furneaux, 1997).

Recent developments in non-human primate neurophysiology help us understand how some of this learning might come about. A variety of experiments have shown that neural mechanisms of reward are integral to the generation of saccadic eye movements. For example, cells in the caudate nucleus of the basal ganglia reflect both the target of an upcoming saccade and the reward expected after making the movement (Hikosaka, Takikawa, & Kawagoe, 2000; Watanabe, Lauwereyns, & Hikosaka, 2003), and cells in other regions of the basal ganglia behave in ways predicted by models of reinforcement learning (Montague, Hyman, & Cohen, 2004; Schultz & Dickinson, 2000). Cortical areas involved in saccade target selection, such as the lateral intra-parietal cortex, exhibit sensitivity both to reward magnitude and to reward

probability (Dorris & Glimcher, 2004; Glimcher, 2003; Platt & Glimcher, 1999; Sugrue, Corrado, & Newsome, 2004, 2005). Cells in the supplementary eye fields also signal the animal's expectation of reward and monitor the outcome of saccade (Stuphorn, Taylor, & Schall, 2000). Since some kind of sensitivity to reinforcement is necessary for learning, and the saccadic circuitry demonstrates such sensitivity, these and other findings reveal a potential neural substrate for learning where to allocate gaze, and learning which objects may be relevant for a task (Bichot & Schall, 1999).

However, unlike physiology experiments in monkeys, human eye movements in daily life do not typically result in receiving a drop of juice. Nevertheless, the above studies suggest that the visual system may generate some kind of internal reward signal after having located relevant information through gaze shifts, and these reward signals may drive the saccadic system to increase the frequency of saccades to objects associated with a positive outcome in the task. Such learning may be contributing to the improved performance in a variety of tasks, including not only visual search, but also more complex tasks extended in time (Hayhoe & Ballard, 2005).

A number of experiments have demonstrated that the visual information detected in a scene, or where subjects choose to attend, depends on their estimates of where this information is likely to be located (Chun & Nakayama, 2000; Eckstein, Pham, & Shimozaki, 2004; Eckstein, Shimozaki, & Abbey, 2002; Landman, Spekrijse, & Lamme, 2003; Posner, 1980; Schmidt, Vogel, Woodman, & Luck, 2002). However, subjects' estimates of statistical structure are often assessed using images of natural scenes, with which observers have a lifetime of exposure (Chen & Zelinsky, 2006; Eckstein, Drescher, & Shimozaki, 2006; Hidalgo-Sotelo, Oliva, & Torralba, 2005; Torralba, Oliva, Castelano, & Henderson, 2006), or derived from experimenters' explicit instructions (Posner, 1980; Palmer, Ames, & Lindsey, 1993). Thus, while knowledge of scenes, or scene structure, has been demonstrated in the above examples, there are few direct demonstrations that humans spontaneously learn where to attend or to direct gaze, the time course of this learning, or what visual cues might be used to direct this learning. However, improved visual search performance has been shown for targets more likely to appear in certain locations (Geng & Behrmann, 2005), or in repeated spatial context (Chun & Jiang, 1998), and the decreased reaction times in those tasks may be facilitated by shifts in gaze behavior (Peterson & Kramer, 2001; Walthew & Gilchrist, 2006). The goal of the present experiments was to develop a paradigm to assess whether observers could learn which objects in a scene would likely be relevant for a task and, consequently, influence the distribution of gaze and subsequent perceptual decisions. The paradigm that we developed was a modified change detection task in which we manipulated the probability that each object would undergo a change on each trial. Although the primary

motivation was to study eye movements, the consequences of our findings also have implications for understanding of how changes are detected and the phenomenon of "change blindness."

Poor performance in change detection tasks has traditionally been interpreted as a result of both the limitations of visual attention and the low capacity of working memory (Rensink, 2000, 2001; Simons, 2000). In addition to these constraints, it is also possible that observers are distributing their gaze sub-optimally. This seems likely, given that the objects undergoing changes in demonstrations of change blindness are typically arbitrary and unlikely to occur in the real world (e.g. the position of a building). Since observers are more sensitive to changes when subjects attend to the location of the change (Schmidt et al., 2002), changes may be more quickly detected if subjects knew where to direct gaze. If some objects are more likely to change than others, subjects may learn to direct gaze preferentially to those objects and maintain this information across a disruption. In the natural world, this would correspond to observers preferentially directing gaze to objects that are likely to change, such as pedestrians and cars, rather than objects that are likely to remain more stable, such as a building.

While the idea that observers' estimates of object changes might influence performance in change detection is not new, this idea has not been thoroughly pursued. Beck, Angelone, and Levin (2004) sought to determine if prior knowledge of the probability of change would influence the detection of visual changes using images of natural scenes. Subjects were more likely to detect probable changes (e.g. the opening and closing of window blinds) than improbable changes (e.g. a door with a window being replaced by a solid door). By using images of natural scenes, however, subjects were presumably relying on their long-term familiarity with the changing objects rather than directly manipulating it. Moreover, the probability an object would change was broadly categorized as either "probable" or "improbable," making it difficult to assess subjects' sensitivity to these statistics for strategic deployment of attention. Finally, since eye movements were not recorded, it is impossible to compare patterns of gaze with behavioral results.

Other evidence that subjects are sensitive to changes in the environment and modify their fixation patterns accordingly has been demonstrated by Droll, Hayhoe, Triesch, and Sullivan (2005) and Jovancevic, Sullivan, and Hayhoe (2006). Droll et al. investigated a task where subjects sorted blocks in a virtual environment and attempted to detect occasional changes in block features. An interesting observation in their experiment was subjects' behavior in subsequent trials after a feature change was successfully detected. On trials following a noticed change, the total time spent fixating the brick was significantly longer on the trial immediately following a detected change, by as much as 400 ms. This effect fell off sharply over the next few trials. This suggests that they

reallocated attentional and gaze resources to the brick on the next trial, immediately following a change. However, since the event was not immediately repeated, they quickly reverted to the prior strategy of infrequent fixations. Jovancevic et al. (2006) observed a similar effect in a virtual walking environment. Following detection of a potential collision with a pedestrian, subjects increased the duration of their fixations on all pedestrians by over 300 ms. These observations of performance show that subjects are sensitive to the probability of events in the scene that are of significance to them, and rapidly adjust gaze strategies in a manner which increase attention towards objects that were recently relevant, and may be expected to be relevant in the future.

As demonstrated in the above experiments, monitoring gaze during change detection tasks is critical to understanding both how observers detect object changes, because successful detection of a change may require fixation patterns similar to those required in visual search. Similar to directing gaze to a target to detect its presence, directing gaze to changing objects may often be necessary, as observers are often insensitive to change in the peripheral retina (O'Regan, Deubel, Clark, & Rensink, 2000), and change detection is more likely following fixations to the changing object (Hollingworth & Henderson, 2002). Rather than searching for a target with known features in a static image, the targets of the search during change detection tasks are objects that might be undergoing a change.

We ask here whether observers learn which objects are most likely to change and allocate gaze accordingly. The present series of experiments were designed to test the hypothesis that observer's fixation patterns can be manipulated by controlling the probability of changes in the

scene. Do observers learn to preferentially direct their gaze towards objects more likely to change, compared to objects expected to remain in a stable state? In addition, does a bias in fixation behavior accompany a differential performance in change detection between objects with different change probability?

## Methods

### Stimuli and trial structure

Stimuli were generated and presented on a Dell PC computer with an 18.5-inch monitor. Experiments were conducted using in-house software written in Matlab, using the Psychophysics (Brainard, 1997; Pelli, 1997) and Eyelink (Cornelissen, Peters, & Palmer, 2002) toolbox extensions. Head position was not fixed, and subjects sat a comfortable distance from the display (approximately 55 cm).

Figure 1 shows the trial structure. To initiate a trial, participants fixated a small white annulus on a gray background and pressed the space bar to initiate each trial. As each trial began, the fixation dot disappeared and participants were briefly presented with a gray blank screen (~2 s). Next, participants were presented with a sequence of "scenes." Each scene was presented for 400 ms and consisted of an array of eight objects. Each object was a white rectangle with a black abstract shape chosen from a set developed by Fiser and Aslin (2001). The size of each object was 1.7 cm × 2.0 cm and subtended approximately 2.1 deg visual angle. Each object

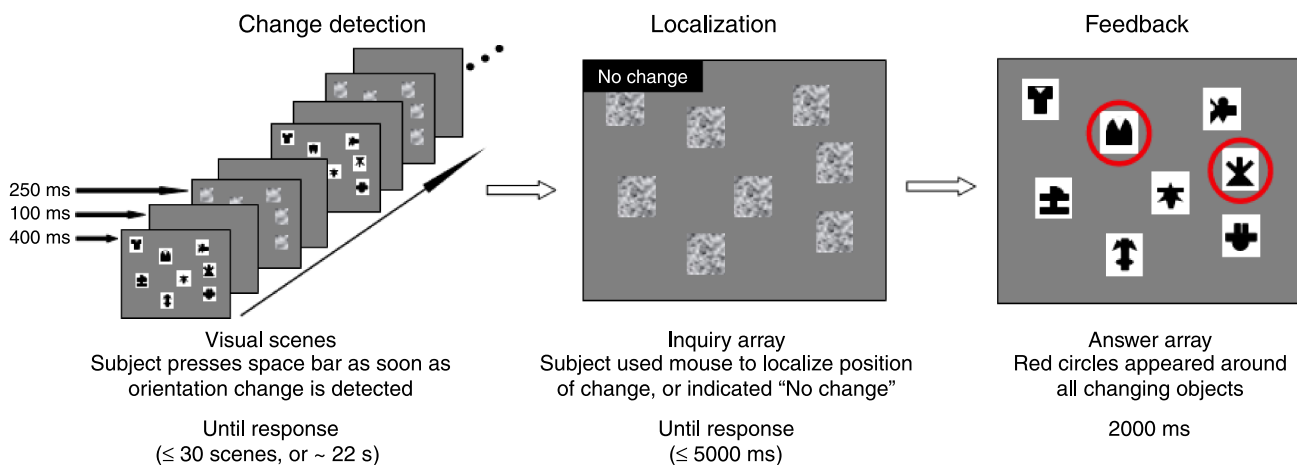


Figure 1. Trial structure for all experiments. Participants were presented a series of up to 30 scenes in which none, one, or multiple objects made 180 deg changes in orientation. Participants pressed the space bar as soon as they detected a change and were immediately presented with an *inquiry array*, in which squares of white noise were located at the previous position of the objects (similar to the mask). The *inquiry array* remained on the screen (for a maximum of 5 s) until participants used the mouse to localize the position of the changed object, or to select "No Change" if they thought no change was present. The *answer array* was then presented for 2 s to provide feedback. The answer array showed the initial scene in that trial, with red circles around each object that was changing on that trial.

was randomly assigned to a position in  $5 \times 4$  invisible array and was jittered with a range of  $\pm 3.0$  deg so that objects appeared to be irregularly scattered on the screen.

In any given scene, each object was in either one of two orientations, rotated  $180^\circ$  up or down. Each scene presentation was followed by a blank gray screen for 100 ms, then by a 250-ms mask, and then by another 100-ms blank screen. The mask was an array of squares containing white noise, in the same positions as the objects in the scene. Participants were then presented with the next scene, in which the positions and shapes of the objects were exactly the same as in the first scene, except that either zero, one, or more of the objects had a different orientation. Participants were instructed to press the space bar as soon as they detected a change in any object, or whenever they were confident that no change was occurring. Trials automatically terminated if the participant did not press the space bar after 30 scene presentations.

After pressing the space bar, the stimulus sequence was immediately interrupted and participants were presented with an *inquiry array*. The *inquiry array* was the same as the mask: an array of white noise squares at the locations of the objects in the scene. The *inquiry array* was presented for a maximum of 5 s or until participants used the mouse to click on the location of the object they first detected as having changed. To accommodate the small number of trials in which there was no change (6% to 17% of trials), participants also had the option of clicking a “No Change” icon in the upper left of the inquiry array to indicate their response that no object was changing on that trial. Following the inquiry array, an *answer array* was presented for 2 s. The *answer array* was a display of the first scene presented in the trial, but with red circles around each of the objects that were changing on that trial. If no change had occurred, the *answer array* was still presented, but without any red circles. Displaying the red circles around each of the changing objects allowed participants to evaluate their own performance and also provided information about which objects were changing on that trial, regardless of participants’ detection of the changes (Figure 1).

Participants performed three practice trials before the data were recorded; on these practice trials, each object was equally likely to change. The experiment consisted of four blocks of trials; each block was generated independently of the others but each within the same statistical constraints. Each block contained 30 trials in Experiment 1 (120 trials total) and 50 trials in Experiment 2 (200 trials total). Participants took a short break halfway through the experiment and were allowed to resume whenever they were ready. The entire experiment took 45–60 minutes for each participant, including calibrating the eye tracker. Participants were recruited either from posters on campus or undergraduate courses in the Brain and Cognitive Sciences Department at the University of Rochester. Participants were paid \$10/hr. All experiments

were tested in accordance to regulations required by the University of Rochester Human Subjects Review Board.

## Monitoring gaze

Gaze was monitored by an Eyelink II eye tracker (SR Research), using a separate PC computer. Both eyes were tracked, whenever possible, and each eye position was recorded at 250 Hz. The experimenter calibrated the eye tracker by having participants fixate nine dots at various locations equally distributed across the screen. For each eye, calibration was considered adequate if the average error was less than one degree and did not exceed 2.0 deg at any of the nine calibration positions. While there was no physical restraint against head movements, participants were encouraged to keep their heads reasonably still. Minor head movements were easily accommodated by the eye tracker. After initiating each trial by fixating the central dot and pressing the space bar, the eye tracker performed a drift correction when necessary. Participants were recalibrated during the short break halfway through the experiment. Objects were considered to be the target of a fixation if the direction of gaze was within the grid square in which the object was positioned, regardless of the random location jitter for each object.

## Experiment 1: Probability of object changes

Experiment 1 was designed to investigate whether observers could learn the probability with which objects underwent a change, and whether this learning influenced sensitivity to object changes and gaze control. Each of eight objects was assigned a fixed frequency of changing across all trials: 0.6, 0.3, 0.1, or 0. Thus, within each block of 30 trials, each object changed 18, 9, 3, or 0 times (and thus, among the eight objects, there was an average number of 2 changes per trial). Subjects were *not informed* that some objects were more likely to undergo a change. Specific trial numbers in which an object would undergo a change were randomly assigned within each of the four blocks of thirty trials. Frequently changing objects had a naturally high probability of undergoing a change on any particular trial. We refer to objects with a high probability of changing as objects with “low stability,” and those with a low probability of changing as objects with “high stability.” Note that because the probability of each object changing is independent from the probability of any other object changing, more than one object could be changing during any particular trial. Despite these multiple changes, we expect that observers’ attention would not be distracted by the possibility of multiple targets from which to select their response, as it

has been noted that observers are nearly incapable of perceiving more than one change at a time (“change simultagnosia”) (Rensink, 2002). Participants were instructed to terminate the trial as soon as they first localized a changing object. Nine subjects performed 120 trials.

## Results and discussion

Subjects accurately localized the position of a changing object in over 95% of change trials, indicating satisfactory performance of the task. A small minority of trials included no changing objects (6.35%). Figure 2 plots the measures of behavioral performance and fixation behavior. The first measure of behavioral performance was the reaction time between the start of the first scene and the time at which subjects pressed the space bar, indicating detection of an object change. In the first block of thirty trials, subjects detected changes to all objects with about the same reaction time, regardless of their probability of changing (Figure 2A) (ANOVA,  $F(2, 8) < 1$ ). However, by the fourth and final block of thirty trials, the pattern of reaction time diverged, with subjects becoming slowest at detecting infrequent changes, and slightly faster at detecting changes that occurred more frequently. (The high variability of reaction time throughout the course of the experiment for high-stability objects is due to the small

number of trials in which those objects changed, and the number of trials in which these changes were reported.) Throughout the course of the experiment, there was a significant interaction in reaction time between object group and trial block (two-way ANOVA,  $F(2, 24) = 4.51$ ,  $p < 0.01$ ).

A second measure of performance was the frequency with which changing objects were detected. Because some objects underwent a change more frequently than others, direct comparisons between the detection rate for each object would not necessarily be meaningful, as these differences may naturally arise as a consequence of the different rate of changes for each object. Thus, in order to effectively compare the selection rates of objects that changed with varying frequencies, it was necessary to consider the number of trials in which each object was reported as having changed, as a percent of the total number of trials in which the object underwent a change. Note that if the selection of changed objects was randomly distributed in each trial, then the detection rate for all changing objects would be equal, regardless of the frequency with which the objects underwent a change within a trial block. The frequency with which observers detected changing objects diverged as the experiment progressed (Figure 2B). Initial rates of change detection were equal for all objects, but subsequent detection rates increased for frequently changing objects and decreased for more stable objects. Detection rates for each object

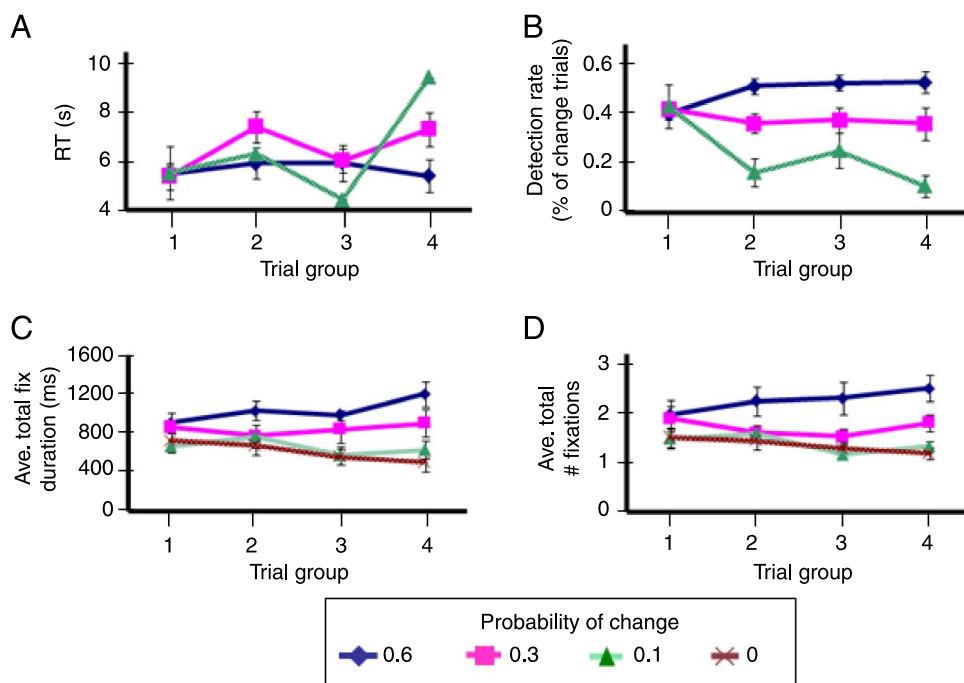


Figure 2. Results from Experiment 1. Frequently changing objects were detected more quickly (A) and more often (B) than objects that rarely underwent a change. Frequently changing objects were also fixated more frequently (C) and for longer total duration (D) than objects with lower probabilities of changing. Fixation data is taken only from trials in which the target object was not undergoing a change (see Methods). Each trial block represents thirty trials and data are collapsed across objects with similar probability of change (see legend).  $N = 9$ .

group across the entire experiment were significantly different (ANOVA,  $F(2, 8) = 11.33, p < 0.001$ ), and there was a significant interaction between object group and trial block (two-way ANOVA,  $F(2, 24) = 3.75, p < 0.005$ ). The increased tendency to report changes to objects with low stability, and the shorter reaction time to report these changes, suggests that observers' attention was directed towards objects that frequently underwent a change.

We next examined whether this change in performance corresponded to a change in fixation behavior. To do this we examined only those trials where a particular object was *not* undergoing a change. This was to avoid a possible confound between the detection of a change and the direction of gaze. A changing object may attract a fixation as a consequence of detection of the change in the periphery, or if a subject chooses to confirm the change by maintaining fixation on the object across successive presentations of the scene. In this case, more frequent fixations to a changing object, or longer fixation durations, may simply be an artifact of the detection the change, and not necessarily reflect a change in learned scene statistics or gaze strategy. Consequently, to ensure that fixation measures were not confounded with the task of detecting visual changes, all fixation data reported are from trials in which the object was not undergoing a change.

Figures 2C and 2D plot average total fixation duration and fixation frequency as a function of trial block. In the first block of thirty trials, all objects were fixated with equal frequencies and for similar durations. However, as the experiment progressed, the pattern of fixation behavior diverged. Observers made a greater number of fixations, and for longer total duration, to objects that were more likely to undergo a change. Throughout the entire experiment, both the number of fixations and the total fixation duration were significantly different for each object group (ANOVA, number of fixations:  $F(3, 8) = 17.96, p < 0.001$ ; fix duration:  $F(3, 8) = 12.44, p < 0.001$ ), and there was a significant interaction between change probability and trial block (two-way ANOVA, number of fixations:  $F(3, 24) = 3.25, p < 0.005$ ; fix duration:  $F(3, 24) = 5.55, p < 0.005$ ).

Note that the patterns of change detection performance in Figures 2A and 2B correspond to a similar pattern of results in the fixation data in Figures 2C and 2D. At the beginning of the experiment, both reaction time and percentage detection are roughly equally for all three categories of changing objects, and the distribution of fixations is most similar among the objects. By the final block of trials, subjects are quickest to detect objects with low stability, are more likely to report their change, and direct gaze to these objects more frequently and for longer duration. This pattern of results suggests that after a relatively short exposure to a series of novel scenes, observers' sensitivity to object changes and their direction of gaze is modulated by the statistical structure underlying these scenes.

## Experiment 2: State-contingent probabilities

The results from Experiment 1 provide evidence that observers were learning the frequency of object changes. This learning is reflected in the fixation patterns and performance in the task of change detection. However, it is not clear what properties of the object observers may be using to learn this statistic. Observers focused their attention on the objects that changed most often, even on trials when those objects were stable. Thus, it is likely that observers are learning something about the predictive value of the shape of the object and the likelihood that the shape will undergo a change. However, this leaves open the question of how specific the information was that observers were using to learn this relationship. Might subjects have been more sensitive to learning what object shapes were predictive of an orientation change, because the shape orientation was implicitly relevant in the change detection task? To what degree might observers learn features irrelevant to the change detection task, but nevertheless predictive of a change? For example, while the position of a car at the side of the road may be unlikely to change, a parked car with its blinker on may be predictive of an upcoming change in position.

## Methods

Participants in Experiment 2 performed the same task of detecting changes in object orientation, but with the following difference. Each trial block contained 50 trials, and each object changed on 10 of those trials (20%). Thus, among the eight objects, there was an average of 1.6 changes on each trial, similar to Experiment 1. The white square containing each object was surrounded by a colored border that was either green or purple and measured approximately 0.6 cm wide, subtending approximately 0.6 deg visual angle. Each object was surrounded by a green border on half of the trials and a purple border on the other half. However, objects surrounded by one of the colors consistently indicated an increased probability that that object would undergo a change on that trial. (The particular color, green or purple, associated with increased change probability was counterbalanced across subjects.) Of the 10 changes for each object per block, 9 occurred when the object was surrounded by the predictive color (18% of trials), and the object only changed once when surrounded by the opposite color border (2% of trials). Subjects were *not informed* about these probabilistic manipulations and were again instructed to report the first change they detected. Twelve participants performed 200 trials.

## Results and discussion

Similar to [Experiment 1](#), observers accurately localized changes in a high percentage of trials (above 95%). A minority of trials included no changing objects (16.78%). The larger number of no-change trials is larger than in [Experiment 1](#) (6.35%) as a consequence of the different distribution of probability values for each object. However, on average, the number of object changes in each experiment was equal (1.6 per trial).

[Figure 3](#) plots observers' behavioral performance and distribution of fixations throughout the four blocks of fifty trials. Unlike the results in [Experiment 1](#), our measures of performance revealed virtually no difference in behavior with respect to the probability with which an object was likely to undergo a change. Observers' reaction time ([Figure 3A](#)), percent detected ([Figure 3B](#)), average total fixation duration ([Figure 3C](#)), and average number of fixation ([Figure 3D](#)) all failed to reveal differences between object state across all trial blocks (ANOVA,  $F < 1$  for all comparisons). (For reasons described above, all fixation data were taken from trials in which the objects being fixated were not undergoing a change; see [Experiment 1](#).) It is not clear if observers failed to learn the predictive value of border color in detecting changes in orientation due to the inherent difficulty of such a calculation or its convenience. Possible reasons for this failure to learn are addressed in the [General discussion](#).

It is also worth noting that due to the differences in statistical structure between [Experiments 1](#) and [2](#), it is not necessarily appropriate to compare performance between these two conditions. For example, the slightly larger number of no-change trials in [Experiment 2](#) may have contributed to the slight overall increase in reaction time.

## General discussion

The present experiments were designed to test whether gaze control and detection of object changes are shaped by the learned likelihood of object changes. To do this, we examined performance in a change detection task where some objects were more likely to change orientation than others. These changes were more likely to occur for certain shapes ([Experiment 1](#)) or were more likely when the shapes were surrounded by a different color border ([Experiment 2](#)). In neither experiment were subjects informed of our manipulations of change or color.

In [Experiment 1](#), observers were quicker to detect changes to objects with low stability and were more likely to report these changes. Conversely, observers were slower at detecting changes to more stable objects and reported a change less often. This pattern of behavior accompanied a shift in subjects' distribution of gaze

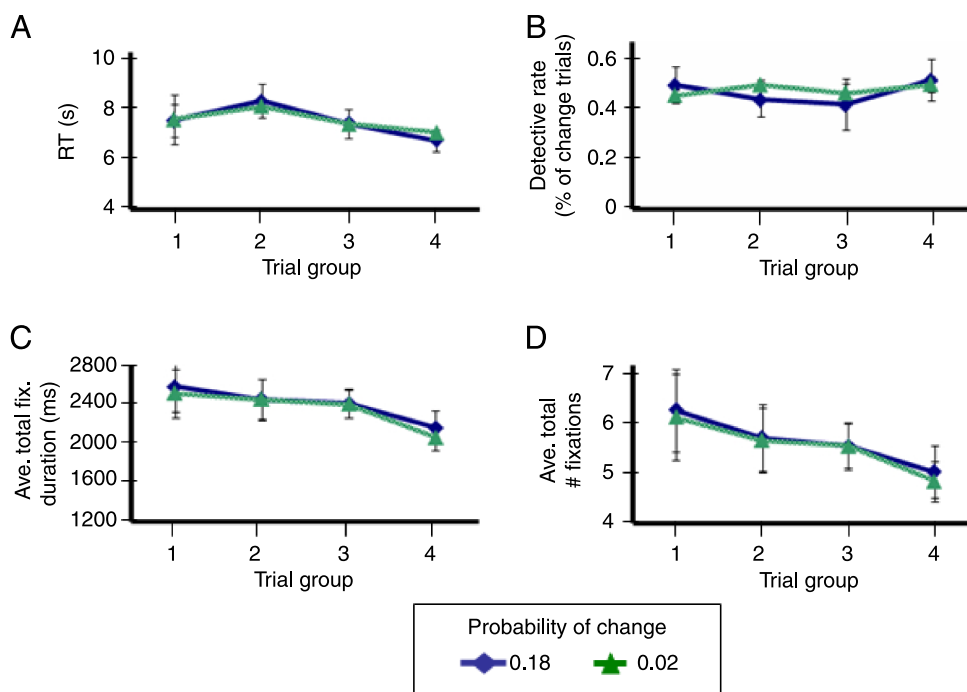


Figure 3. Results from [Experiment 2](#). Despite the fact that the color of the surrounding border for each object was predictive of whether the object would undergo a change on that trial, subjects did not show differential behavior for this attribute, as assessed by measures of reaction time (A), proportion of changes detected (B), average fixation duration (C), or number of fixations to each object (D). Each trial block represents fifty trials (200 total).  $N = 12$ .

towards objects, even when those objects were not undergoing a change. Objects with low stability were more likely to be fixated, and for longer duration, than objects less likely to undergo a change.

However, in [Experiment 2](#), subjects did not appear to learn conditional probabilities of change over the course of the experiment. The presence of a colored border, predictive of a change in orientation, did not influence the detection of changes, nor did it influence the distribution of gaze. The results from these experiments have implications for mechanisms gaze control in scenes, strategies for updating internal representations of dynamic environments, and understanding the underlying cause of change blindness.

## Learning gaze control

The current experiment shows that observers rapidly learn the probability of a change and use it to direct gaze. As described in the [Introduction](#), recent work on the neurophysiology of eye movements has suggested that neural activity in many areas provides a substrate for learning. Cortical signals encoding the learned probability of certain events, or the utility for making an eye movement to particular stimuli, manifest themselves through increased neural response as well as behavioral measures, such as the likelihood of a saccade to a target (Glimcher, 2003). In a neurophysiological experiment, reward can be directly manipulated. However, in human eye movement behavior it is not entirely clear what constitutes reward. In this paper, we assume that finding the desired information, in this case a changing object, provides some kind of secondary reward. To the extent that locating information is important for the organism this seems like a reasonable assumption. Thus, by manipulating probability that an object will be changing, we are effectively manipulating reward probability. Similar manipulation of the probability that a stimulus will be rewarded, or the probability that a stimulus will be presented, has been shown to affect responses of cells both in the lateral intraparietal cortex and in buildup neurons in the intermediate layers of the superior colliculus, the mid-brain saccade generator (Basso & Wurtz, 1998; Platt & Glimcher, 1999). This forms a parallel with psychophysical observations showing that saccade reaction time is similarly influenced by stimulus probability (He & Kowler, 1989). The present experiment demonstrates the importance of stimulus probability in a more complex situation, where subjects learn that particular objects are more likely to change than others. Reaction time and probability of reporting the change as well as number of fixations and fixation duration were all modified within 60 trials. Because subjects were not informed about the different probability of changes across objects, this learning may reflect an intrinsic sensitivity to the probabilistic structure of the environment. Gaze

patterns observed during more complex tasks, or when viewing images of natural scenes, may thus be similarly shaped by the subjects' experience with similar scenes encountered over the course of a lifetime (Eckstein et al., 2006).

A natural question is whether the observed learning in [Experiment 1](#) is a signature of short-term or long-term memory representations. In laboratory settings, observers have shown remarkable capacity and retention for previously shown images of natural scenes (Nickerson, 1968; Standing, 1973). However, it is not clear if memory for statistical structure is similarly robust. However, given that performance differences were continuing to diverge throughout the present experimental sessions of 60 minutes, this suggests that this learning is potentially long-term (Fiser & Aslin, 2001, 2002). However, because we did not explicitly test for the duration of these representations, this may best be addressed in future research.

## Implications for change blindness

Change detection paradigms can be used to test many cognitive processes, including the capacity of working memory (Wheeler & Treisman, 2002), mechanisms of comparing internal information (Simons & Rensink, 2005), and the use of visual information throughout the course of extended tasks (Droll & Hayhoe, [in press](#)). Similar to the use of change detection during extended tasks, in the present paradigm, we use sensitivity to changes in conjunction with fixation behavior to assess what visual information in the scene is selected for encoding and storage. Like other change detection paradigms, the changes in the present experiment were trivially simple to detect, once observers knew where to look. Noticing an object's change in orientation between two scenes is well within the capacity of working memory. The bottleneck in performance was in developing a strategy of where to direct gaze in order to acquire and compare the relevant information. By framing this change detection task as a challenge of gaze control, our paradigm, similar to many other change detection experiments, is akin to a visual search task, in which the target of the search is the object considered most likely to undergo a change.

The observed spontaneous and rapid learning of the probability of object changes and its influence on change detection and gaze control have implications for understanding the phenomenon of change blindness. Performance in change detection experiments has traditionally been interpreted as evidence for the visual system's limited capacity to represent detailed scene information and the maintenance of only sparse internal representations (O'Regan, 1992). While this generalization may reflect some of the fundamental limitations of vision, it does not explain how observers can successfully navigate



within real-world dynamic visual environments in which object changes are frequent, despite such apparent blindness to change in laboratory demonstrations. However, one way in which to reconcile our results with previous literature is if estimates of the probability of object changes are included in observers' internal representation of scene context (Rensink, 2000). This would suggest that demonstrations of change blindness using images of natural scenes may stem from the mismatch between the observer's expectations and the particular changes employed in the experiments. While an observer's expectations are learned over the course of a lifetime, the object changes in the laboratory are often quite arbitrary, so observers may simply not know where to look.

Previous accounts of change blindness have sometimes addressed the knowledge of a scene that an observer has before a change occurs. Rensink, O'Regan, and Clark (1997) found that changes to objects of central interest in an image of a natural scene were detected faster than objects of marginal interest. This central interest advantage was interpreted as a consequence of visual attention being drawn to objects as a function of their semantic contribution to the scene. This interpretation is broadly consistent with the idea that subjects preferentially direct their gaze towards objects with which they have found to be important for a task. However, such demonstrations of change blindness in natural scenes do not take into account subjects' prior knowledge of which objects might naturally be expected to change, and the changes that occur are seemingly arbitrary (but see Beck et al., 2004). For example, in the Rensink et al. experiment, changes to objects categorized as being of marginal interest were sometimes changes that are arguably unlikely to occur (e.g. the position of a handrail), whereas changes to objects considered to be of central interest were sometimes changes one would consider likely to occur (e.g. the position of a helicopter in flight) (Rensink et al., 1997, Figure 2). Thus, it is not clear how expectations of object changes may have interacted with, or contributed to, the semantic meaning of the scene. It is possible that observers may have been using expectations of likely object changes, rather than a higher level understanding of scene gist, to guide attention and gaze. Our present use of artificial "objects" with which observers have no prior experience sidesteps any concern of semantic object meaning, or scene gist, and may thus reflect a general learning mechanism that is applied to novel environments in which the statistics are initially unfamiliar.

Another experiment using artificial scenes and demonstrating learning during change detection is Jiang and Song (2005). Those researchers have shown that, similar to the contextual cueing effect in visual search (Chun & Jiang, 1998), measures of decreased reaction time and increased percent correct reveal learning of repeated spatial context in change detection tasks. However, because eye movements were not monitored in that study, it is not clear how observers' performance was

influenced by their deployment of gaze. In the present study, while change detection was the primary task, the distribution of gaze on non-changing objects suggests that observers were performing a search for the objects they expected to be undergoing a change. Thus, because the task of change detection was intertwined with the secondary task of visual search, the contribution of learning from each of these tasks cannot be easily disassociated.

## Limits of learning

While observers were quick to learn and to exploit the probability of object changes (Experiment 1), there was no evidence of learning which color state was predictive of a change in orientation (Experiment 2). It is not clear if observers are less sensitive, or perhaps incapable, of learning such state-contingent features whose state predict changes in other object features. A notable difference between the two paradigms is that in Experiment 1, the shape of an object in a single frame was both predictive of whether a change would occur at that location, as well as being relevant to the task of detecting the change in orientation. This double role of object shape may have made it easier for subjects to associate object shape with the probability of change. In contrast, in Experiment 2, border color was only incidental to the task of detecting a change in orientation, despite its significant power in predicting this change. It is not clear whether observers did not encode color information in the first place because of its irrelevance in the primary task of change detection (Hayhoe, Bensinger, & Ballard, 1998), or whether they encoded color information, but failed to associate it with the likelihood of an orientation change.

If subjects had encoded color, there are still significant constraints on observers' ability to learn the relationship with an orientation change. These constraints may include the degree of structure in the stimuli, observers' intrinsic capability of learning these statistics, the utility of applying this knowledge to improve performance, and the convenience with which the learner can compute these relationships (Hunt & Aslin, 2001). While observers have demonstrated learning complex relationships of scene structure in other visual tasks (Fiser & Aslin, 2001, 2002), the particular combination of task and statistics in Experiment 2 may have been too difficult for observers to learn and exploit within the short experimental session. Michel and Jacobs (2007) have also suggested that learning contingencies between cues and their predictive variables may occur only when observers consider these relationships to be potentially dependent. Thus, the predictive relationship between color and orientation change may not have been learned either because we did not inform subjects of this possible association, or this relationship may not have been spontaneously considered by the observers during the task.

## Conclusions

The present experiments have used a change detection paradigm to demonstrate that gaze control and change detection may be influenced by the leaned likelihood of object changes in visual scenes. Thus, change blindness may be considered in part, as a consequence of not knowing where to look and what to attend.

## Acknowledgments

This work was supported by National Institutes of Health/Public Health Service Research Grants EY07125, 815 EY05729, and RR09283, and by IC Postdoctoral Research Fellowship Program Grant 8-444069-23149 (icpostdoc.org).

Commercial relationships: none.

Corresponding author: Jason A. Droll.

Email: droll@psych.ucsb.edu.

Address: Department of Psychology, University of California Santa Barbara, Santa Barbara, CA 93106, USA.

## References

- Basso, M. A., & Wurtz, R. H. (1998). Modulation of neuronal activity in superior colliculus by changes in target probability. *Journal of Neuroscience*, *18*, 7519–7534. [[PubMed](#)] [[Article](#)]
- Beck, M. R., Angelone, B. L., & Levin, D. T. (2004). Knowledge about the probability of change affects change detection performance. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 778–791. [[PubMed](#)]
- Bichot, N. P., & Schall, J. D. (1999). Effects of similarity and history on neural mechanisms of visual selection. *Nature Neuroscience*, *2*, 549–554. [[PubMed](#)] [[Article](#)]
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. [[PubMed](#)]
- Chen, X., & Zelinsky, G. J. (2006). Real-world visual search is dominated by top-down guidance. *Vision Research*, *46*, 4118–4133. [[PubMed](#)]
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71. [[PubMed](#)]
- Chun, M. M., & Nakayama, K. (2000). On the functional role of implicit visual memory for the adaptive deployment of attention across scenes. *Visual Cognition*, *7*, 65–81.
- Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The EyeLink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments & Computers*, *34*, 613–617. [[PubMed](#)] [[Article](#)]
- Dorris, M. C., & Glimcher, P. W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron*, *44*, 365–378. [[PubMed](#)] [[Article](#)]
- Droll, J. A., & Hayhoe, M. (in press). Trade-offs between gaze and working memory use. *Journal of Experimental Psychology: Human Perception and Performance*.
- Droll, J. A., Hayhoe, M. M., Triesch, J., & Sullivan, B. T., (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1416–1438. [[PubMed](#)]
- Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting and Bayesian priors. *Psychological Science*, *17*, 973–980. [[PubMed](#)]
- Eckstein, M. P., Pham, B. T., & Shimozaki, S. S. (2004). The footprints of visual attention during search with 100% valid and 100% invalid cues. *Vision Research*, *44*, 1193–1207. [[PubMed](#)]
- Eckstein, M. P., Shimozaki, S. S., & Abbey, C. K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision*, *2*(1):3, 25–45, <http://journalofvision.org/2/1/3/>, doi:10.1167/2.1.3. [[PubMed](#)] [[Article](#)]
- Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*, 499–504. [[PubMed](#)]
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 458–467. [[PubMed](#)]
- Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics*, *67*, 1252–1268. [[PubMed](#)]
- Gibson, B. S., & Jiang, Y. (1998). Surprise! An unexpected color singleton does not capture attention in visual search. *Psychological Science*, *9*, 176–182.
- Glimcher, P. W. (2003). The neurobiology of visual-saccadic decision making. *Annual Review of Neuroscience*, *26*, 133–179. [[PubMed](#)]
- Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*, 188–194. [[PubMed](#)]
- Hayhoe, M. M. (2000). Vision using routines: A functional account of vision. *Visual Cognition*, *7*, 43–64.

- Hayhoe, M. M., Bensinger, D. G., & Ballard, D. H. (1998). Task constraints in visual working memory. *Vision Research*, *38*, 125–137. [PubMed]
- Hayhoe, M. M., Droll, J. A., & Mennie, N. (2007). Learning where to look. In R. P. G. van Gompel, M. H. Fisher, & W. S. Murray (Eds.), *Eye movements: A window on mind and brain* (pp. 641–660). Oxford: Elsevier.
- He, P. Y., & Kowler, E. (1989). The role of location probability in the programming of saccades: Implications for “center-of-gravity” tendencies. *Vision Research*, *29*, 1165–1181. [PubMed]
- Hidalgo-Sotelo, B., Oliva, A., & Torralba, A. (2005). Human learning of contextual priors for object search: Where does the time go? In *Proceedings of the 3rd Workshop on Attention and Performance in Computer Vision, Int. CVPR*. Washington, DC: IEEE Computer Society.
- Hikosaka, O., Takikawa, Y., & Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological Review*, *80*, 953–978. [PubMed] [Article]
- Hollingworth, A., & Henderson, J. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 113–136.
- Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General*, *130*, 658–680. [PubMed]
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506. [PubMed]
- Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews, Neuroscience*, *2*, 194–203. [PubMed]
- Jiang, Y., & Song, J. H. (2005). Spatial context learning in visual search and change detection. *Perception & Psychophysics*, *67*, 1128–1139. [PubMed]
- Jovancevic, J., Sullivan, B., & Hayhoe, M. (2006). Control of attention and gaze in complex environments. *Journal of Vision*, *6*(12):9, 1431–1450, <http://journalofvision.org/6/12/9/>, doi:10.1167/6.12.9. [PubMed] [Article]
- Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, *28*, 1311–1328. [PubMed]
- Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *352*, 1231–1239. [PubMed] [Article]
- Landman, R., Spekreijse, H., & Lamme, V. A. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, *43*, 149–164. [PubMed]
- Michel, M. M., & Jacobs, R. A. (2007). Parameter learning but not structure learning: A Bayesian network model of constraints on early perceptual learning. *Journal of Vision*, *7*(1):4, 1–18, <http://journalofvision.org/7/1/4/>, doi:10.1167/7.1.4. [PubMed] [Article]
- Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, *431*, 760–767. [PubMed]
- Nickerson, R. S. (1968). A note on long-term recognition memory for pictorial material. *Psychological Science*, *11*, 58.
- O’Regan, J. K. (1992). Solving the “real” mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, *46*, 461–488. [PubMed]
- O’Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, *7*, 191–212.
- Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 108–130. [PubMed]
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. [PubMed]
- Peterson, M. S., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics*, *63*, 1239–1249. [PubMed] [Article]
- Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*, 233–238. [PubMed]
- Posner, M. (1980). The orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25. [PubMed]
- Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition*, *7*, 17–42.
- Rensink, R. A. (2001). Change blindness: Implications for the nature of visual attention. In L. R. J. M. Harris (Ed.), *Vision and attention* (pp. 169–188). New York: Springer.
- Rensink, R. A. (2002). Failure to see more than one change at a time [Abstract]. *Journal of Vision*, *2*(7):245, 245, <http://journalofvision.org/2/7/245/>, doi:10.1167/2.7.245.

- Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*, 368–373.
- Schmidt, B. K., Vogel, E. K., Woodman, G. F., & Luck, S. J. (2002). Voluntary and automatic attentional control of visual working memory. *Perception & Psychophysics*, *64*, 754–763. [[PubMed](#)] [[Article](#)]
- Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience*, *23*, 473–500. [[PubMed](#)]
- Simons, D. (2000). Change blindness and visual memory. *Visual Cognition*, *7*, 1–15.
- Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, *9*, 16–20. [[PubMed](#)]
- Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology*, *25*, 207–222. [[PubMed](#)]
- Stuphorn, V., Taylor, T. L., & Schall, J. D. (2000). Performance monitoring by the supplementary eye field. *Nature*, *408*, 857–860. [[PubMed](#)]
- Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science*, *304*, 1782–1787. [[PubMed](#)]
- Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nature Reviews, Neuroscience*, *6*, 363–375. [[PubMed](#)]
- Torralba, A., Oliva, A., Castelhan, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*, 766–86. [[PubMed](#)]
- Walther, C., & Gilchrist, I. D. (2006). Target location probability effects in visual search: An effect of sequential dependencies. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1294–1301. [[PubMed](#)]
- Watanabe, K., Lauwereyns, J., & Hikosaka, O. (2003). Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. *Journal of Neuroscience*, *23*, 10052–10057. [[PubMed](#)] [[Article](#)]
- Wheeler, M. E., & Treisman, A. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*, 48–64. [[PubMed](#)]
- Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 601–621. [[PubMed](#)]
- Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary vs. automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 121–134. [[PubMed](#)]