The insistence of vision: Why do people look at a salient stimulus when it signals target absence?

Christopher W. Myers¹, Wayne D. Gray², and Chris R. Sims³

¹Air Force Research Laboratory, Wright-Patterson, AFB, OH, USA
²Cognitive Science Department, Rensselaer Polytechnic Institute, Troy, NY, USA
³Brain & Cognitive Systems, University of Rochester, Rochester, NY, USA

Researchers and practitioners across many fields would benefit from the ability to predict human search time in complex visual displays. However, a missing element in our ability to predict search time is our ability to quantify the exogenous attraction of visual objects in terms of their impact on search time. The current work represents an initial step in this direction. We present two experiments using a quadrant search task to investigate how exogenous and endogenous factors influence human visual search. In Experiment 1, we measure the oculomotor capture—or the tendency of a stimulus to elicit a saccade—of a salient quadrant under conditions in which the salient quadrant does not predict target location. Despite the irrelevance of quadrant salience, we find that subjects persist in making saccades towards the salient quadrant at above-chance levels. We then present a Bayesian-based ideal performer model that predicts search time and oculomotor capture when the salient quadrant never contains the search target. Experiment 2 tested the predictions of the ideal performer model and revealed human performance to be in close correspondence with the model. We conclude that, in our speeded search task, the influence of an exogenous attractor on saccades can be quantified in terms of search time costs and, when these costs are considered, both search time and search behaviour reflect a boundedly optimal adaptation to the cost structure of the environment.
Visually salient information is typically regarded to exogenously influence saccade processes (i.e., in a bottom-up or data-driven manner; Everling & Fischer, 1998; Findlay, 1982, 1997; Franconeri & Simons, 2003; Franconeri, Simons, & Junge, 2004; Henderson, 2007; Itti, Koch, & Niebur, 1998; Kowler, 1990; Mitchell, Macrae, & Gilchrist, 2002; Pomplun, Reingold, & Shen, 2003; Theeuwes, 2004; Treisman & Gelade, 1980; Wolfe, 1994). Exogenous influences can capture attentional resources (i.e., attentional capture), which may result in saccades towards the attention-capturing location (i.e., oculomotor capture; see Chapter 3 of Findlay & Gilchrist, 2003, for a review). Further, exogenous influences can affect the efficiency of visual search: A target is found more quickly when it is salient than when it is not (Wolfe, 1998) and salient nontargets can result in oculomotor capture that increases visual search time (Theeuwes, Kramer, Hahn, & Irwin, 1998).

Consciously adopted strategies, often referred to as endogenous, also affect the efficiency of visual search (Boot, McCarley, Kramer, & Peterson, 2004; Chernyak & Stark, 2001; Hayhoe & Ballard, 2005; Hayhoe, Shrivastava, Mruczek, & Pelz, 2003; Henderson, 2007; Horowitz & Wolfe, 2001, 2003; Land & Lee, 1994; Land & McLeod, 2000; Peterson, Kramer, Wang, Irwin, & McCarley, 2001; Rayner, 1998; Rosenholtz, 1999; Yarbus, 1967) and can be deliberately adopted through instruction, as in the antisaccade task (Everling & Fischer, 1998; Mitchell et al., 2002), or adopted through task goals, such as making a peanut butter and jelly sandwich, driving a race car, or finding a green X among green Ys and blue Xs. (Hayhoe, 2000; Hayhoe & Ballard, 2005; Hayhoe et al., 2003; Land, Furneaux, & Gilchrist, 2002; Land & Lee, 1994; Land & McLeod, 2000; Land & Tatler, 2001; Yarbus, 1967).

The focus of the current study is on those cases where exogenous features of a task allow an efficient endogenous search strategy, but may also lead to oculomotor capture that opposes the speeded execution of that strategy. Oculomotor capture often occurs in spatial cueing paradigms (Mulckhuyse & Theeuwes, 2010) such as the antisaccade task (Hallett, 1978; see Everling & Fisher, 1998, for a review), in which the abrupt onset of the attentional cue to one side of the display signals that a saccade must occur to the opposite side. Hence, the abrupt onset is an exogenous cue that signals the endogenous strategy of moving the eyes to the opposite side of the display. The fastest responses are for trials in which subjects can inhibit oculomotor capture from the onset and move their eyes to the opposite side of the display. However, although completing the trial by identifying the item is faster when the cued saccade (i.e., prosaccade) is inhibited, prosaccades are
executed faster than endogenously influenced saccades away from the onset (i.e., antisaccades; Godijn & Theeuwes, 2002; Mulckhuyse & Theeuwes, 2010; Nieuwenhuis, Ridderinkhof, de Jong, Kok, & van der Molen, 2000; Wolfe, Alvarez, & Horowitz, 2000). In the antisaccade task, prosaccades are never fully extinguished, reaching asymptote at various proportions of occurrence, from 5% to 7% (Everling & Fischer, 1998) to 11% (Hallett & Adams, 1980), and prosaccade occurrence is especially high in groups with disorders (e.g., schizophrenics ≈ 58%; Everling & Fischer, 1998).

In the current paper we seek to explain the continuation of oculomotor capture when an exogenous cue signals an endogenous strategy of cue avoidance. Intuitively, it would seem that completely avoiding oculomotor capture would speed task performance. We will conclude that this intuition is false and we will provide a quantitative ideal performer model that predicts our results.

The ideal performer model follows in the tradition of ideal observer analyses in the perception and psychophysics community (Geisler, 2003; Macmillan & Creelman, 1991; Najemnik & Geisler, 2008): Within a given task environment, an ideal observer model seeks to uncover the best performance that a cognitive system can be expected to achieve within a set of specified constraints. Whereas ideal observers perform optimal inference over uncertain information, ideal performers must also act on the basis of their inferences. Different actions can have consequences of varying utility, and the outcomes of the actions may also be uncertain. The mathematical framework combining optimal inference over uncertain information and rational decision making is known as Bayesian decision theory (Körding, 2007; North, 1968), and our ideal performer model is derived using this framework.

An ideal performer model answers the why question rather than the how question of behaviour by casting behaviour in a purely mathematical framework that allows researchers to focus on understanding behaviour according to how efficiently it achieves the specified goals for the system. If human behaviour is empirically found to be in close correspondence to the predictions obtained from the model, then that behaviour can be understood as optimally achieving the relevant goals. If behaviour diverges from the predictions of the model, this may lead to insight regarding cognitive limitations or bottlenecks that can account for the discrepancy (Jacobs & Kruschke, 2011). After understanding what goals the behaviour achieves, the next step would be to understand the cognitive or neural mechanisms that would achieve that behaviour.

The theoretical perspective of the work presented is informed by the soft constraints hypothesis (Gray & Fu, 2004; Gray, Sims, Fu, & Schoelles, 2006), which maintains that at the 1/3 to 3 s embodied level of analysis (Ballard, Hayhoe, Pook, & Rao, 1997) sequences of cognitive, perceptual, and motor operations (i.e., interactive routines) are selected to minimize performance
costs measured in time while achieving expected benefits. Cost–benefit tradeoffs provide a soft constraint on the sequential organization of behaviour as they may be overridden by endogenous influences such as training or deliberately adopted strategies. Consequently, if a target can be found with one less eye fixation while maintaining accuracy and increasing search speed, then behaviour will tend to adapt by eliminating the surplus fixation (Myers & Gray, 2010).

The current set of studies introduces a new spatial cueing task, the salient quadrant paradigm, in which the degree of exogenous influence of a salient quadrant and its task relevance can be orthogonally varied. In two experiments the target is always at the centre of one of four quadrants, and the subjects’ task is to locate the target as quickly as possible. The first experiment provides data to quantify the exogenous influence of the salient quadrant to predict oculomotor capture when the salient quadrant is task-irrelevant. Data from Experiment 1 are used to calibrate the ideal performer model to predict oculomotor capture rates when the salient quadrant is task relevant.

In the second experiment the target is never located in the salient quadrant making it task relevant. Hence, an optimal strategy is to never search the salient quadrant thereby reducing the search space from four quadrants to three. There are two questions posed for this experiment: (1) Can participants learn to avoid the salient quadrant? And (2), will avoidance behaviour reach optimal levels given the exogenous influence of the salient quadrant and system noise? The model derived from Experiment 1 data accurately predicts data from Experiment 2 without any adjustments to model parameters and supports our interpretation that the probability of initial saccades reflects an optimal behavioural accommodation of exogenous influences to the cost structure of the search environment.

**EXPERIMENTS**

Two visual search experiments using the salient quadrant paradigm were conducted to investigate exogenous influences on saccades and search efficiency. In the following section we first discuss the general aspects of the salient quadrant paradigm, followed by the salience of the salient quadrant. Experiment 1 is covered next, followed by the derivation of an ideal performer model and predictions, which are then tested in Experiment 2.

**General method**

In the salient quadrant paradigm, subjects are presented with a display consisting of four quadrants, one of which contains the target. The configuration of search displays differs by covarying three factors to produce
one display quadrant that was more salient than its three neighbouring quadrants. For each of the reported experiments there were 100 items on the screen distributed in a $10 \times 10$ grid with a default separation of $1.94^\circ$ of visual angle. (All visual angles were calculated for a fixed viewing distance of 55 cm.) A target item (“L”) was present on every trial and was always located at the centre of one of the four quadrants (see Figure 1). Distractor items (“T”) occupied the other 99 grid positions.

The quadrant saliency manipulation decreased the default separation of the 25 items in the salient quadrant by covarying quadrant density and the apparent quadrant size. Both of the reported experiments used the same levels of quadrant saliency: strong, moderate, and control. The items in the strong saliency quadrant had an interitem distance of $0.54^\circ$, those in the moderate saliency quadrant $0.97^\circ$, and those in the control quadrant used the default interitem distance of $1.94^\circ$ (i.e., all four quadrants had the same degree of item separation on the control trials, see top middle of Figure 1).

In the strong and moderate saliency conditions the salient quadrant was surrounded by an empty frame that separated it from the remaining quadrants as well as from the edges of the search display (see Figure 1). However, the item at the centre of each quadrant occupied the same screen coordinates regardless of the level of quadrant saliency. The space between salient and neighbouring quadrants changed the perceived size of the salient quadrant and provided the quadrant with a clearly defined boundary lacking in the control display.

**Measuring the salience of the salient quadrant**

Salience was defined as the relative difference between an item, or an area, and the surrounding entities or areas. A computational measure of visual salience was used as an operational definition of visual salience (Rosenholtz, 1999, 2001). Similar to other measures (e.g., Itti, 2006; Itti & Koch, 2001; Itti et al., 1998), Rosenholtz’s measure represents each item of a display on the dimensions of colour, orientation, and contrast, and essentially tests for outliers on each of these dimensions. The mean and covariance of each pixel are computed from the distribution of features present in a display. Salience is computed as the Mahalanobis distance between a given pixel and all other pixels (Rosenholtz, 1999, 2001).

The relative quadrant salience for each salience manipulation (i.e., strong, moderate, and control) was computed for each of the four possible quadrant locations. For each quadrant, the saliency values of the pixels in that quadrant were averaged to produce a mean salience score. Salient quadrant locations were then averaged to produce a mean salience score for each of the three levels of salience.

The degree of quadrant salience was the difference between the mean of the manipulated quadrants’ salience score and the other three quadrants’ scores
for each of the strong and moderate conditions. For this comparison, the absolute value of the salience score is not important; rather, the difference between salience scores for the manipulated versus the other three quadrants is the important measure. For the moderate condition, this yields a Rosenholtz difference \( (RΔ = (sq - μ_3)/sq)) \) between the saliency of the salient quadrant \( (sq) \) and the mean of the other three quadrants \( (μ_3) \) of 24%. For the strong condition, the Rosenholtz difference between the salient quadrant and the remaining three was 59%. Thus, Rosenholtz’s visual saliency model predicts that the strong saliency quadrants are more than double the salience of moderate saliency quadrants.

Figure 1. The task environment used in experiments. Each trial began with the fixation control display (left), and was followed by one of three stimulus displays (centre)—control saliency, moderate saliency, and strong saliency. Each stimulus display had a .33 probability of occurring. After locating the target (L), subjects clicked the found button producing the test display. Subjects then clicked on the quadrant (right) in which they found the target in the stimulus display. A new trial began after clicking on the desired quadrant.
Common procedures across experiments

A single trial of the salient quadrant paradigm consisted of a sequence of three screens. Screen 1 was the fixation control screen. This screen was composed of small crosshairs located in the middle of an otherwise empty window (see Figure 1, left). To begin the trial, subjects were instructed to fixate the crosshairs. A remote eyetracking system was used to determine fixations. Once it was detected that a subject had fixated on the crosshairs, the stimulus display replaced the fixation control screen (see Figure 1, centre). The automatic detection of a fixation at the location of the crosshairs served as a control to ensure that all subjects began search from the same location and that subjects remained calibrated to the eyetracker. If the eyetracking system did not register a fixation on the crosshairs after 10 s, the stimulus display would appear. Trials that resulted in inadequate initial fixation control were removed from analyses.

Screen 2, the stimulus display screen (see Figure 1, centre), contained one target (L) and 99 distractors (T). On each trial, the target and each of the distractors were independently and randomly rotated about their axes to one of four possible orientations: $0^\circ$, $90^\circ$, $180^\circ$, and $270^\circ$. There were no visual markings denoting the four quadrants. Indeed, in the control condition the four quadrants appeared as a continuous field of 100 evenly spaced items. For the moderate and strong saliency displays, the salient quadrant was easily discerned, but no overt visual boundary or marking separated the other three quadrants from each other. Each saliency level had a $1/3$ probability of occurring on each trial.

When the stimulus display screen appeared, subjects were required to search for the target and click the “Found” button (see Figure 1, centre) when the target was located. (Note that at the beginning of each trial, the cursor was moved inside of the “Found” button by the experiment software.) The stimulus display screen was replaced by the test display screen (see Figure 1, right) after the participant clicked the “Found” button.

Screen 3, the test display, was blank except for a vertical and horizontal line that divided the screen into four quadrants. Coincident with the appearance of the test display, the mouse pointer was positioned at the lines’ intersection by the experiment software. All subjects were instructed to click on the quadrant where the target was discovered. Once the participant clicked on a quadrant, the pointer was automatically relocated to “Found” and the fixation display reappeared signalling the start of a new trial.

Task relevance of the salient quadrant

The relevance of the salient quadrant to locating the target was manipulated across experiments. In the first experiment, salient quadrants were made
completely task-irrelevant by fully counterbalancing the target and salient quadrant locations. As a result, there was no correlation between the salient quadrant and target location. In Experiment 2, salient quadrants and the target were never colocated. Consequently, the salient quadrant was task relevant—its location could be used to reduce the search space by ruling out a possible target location.

Common apparatus
Across both studies, the experiments were presented on a PowerMac G4 Apple computer running Macintosh OS X 10.2, a 17-inch flat panel display with the resolution set to $1024 \times 768$. Eye data were collected using an Eyegaze® (LC Technologies) eyetracking system that measured gaze point at a 60 Hz rate. A chinrest, used to promote head stabilization, was situated so that subjects’ calibrated eye was 55 cm from the monitor. The task environment was built in-house using Macintosh Common Lisp.

We defined a fixation using a sample-based algorithm. Fixations were operationally defined as maintaining steady gaze within $2^\circ$ visual angle. The minimum requirements for a fixation were six fovea position samples. At a sampling rate of 60 Hz, the minimum fixation duration was 100 ms (following Karsh & Breitenbach, 1983).

By design of the task requirements, the first fixation on the search display was colocated with the crosshairs from the fixation control display. The first saccade was defined as the gaze position samples occurring between the end of the first fixation and the beginning of the second fixation on the search display. The second fixation on the search display was defined as the initial saccade location.

Subjects were instructed to maximize their speed and accuracy in the visual search task. The eye tracker was introduced to subjects as a means of implementing gaze-contingent control to initiate each trial. No mention was made of any other role for the eyetracker.

EXPERIMENT 1
In Experiment 1 the location of the salient quadrant was independent of the target quadrant, making the salient quadrant irrelevant to finding the target. The experiment was conducted to provide evidence that the salient quadrant exerted an exogenous influence on initial saccades, to provide a baseline of initial saccades to the salient quadrant for comparison to Experiment 2, and to provide parameter values for our ideal performer model.
Method

Procedure. Subjects performed four blocks of 48 visual search trials per block. Each quadrant had a 25% chance of containing the target and a 25% chance of being the salient quadrant on each trial. For the salient quadrant, each level (control, medium, and strong) was used equally often. Hence, within each block of 48 trials each combination of salient quadrant location (four) by degree of quadrant saliency (three) by target location (four) was used once. The exact sequence of the 48 combinations was determined at random for each participant.

Subjects. Thirteen undergraduate students from Rensselaer Polytechnic Institute volunteered to participate. All subjects had normal or corrected-to-normal vision. The study lasted approximately 1 hour. Subjects were run individually, and were appropriately compensated for their time.

Results

Results are divided between visual search time and accuracy, effect of the salient quadrant, and saccade reaction time (SRT) analyses. Prior to conducting analyses, trials were excluded for which the gaze-contingent display did not detect an initial fixation on the fixation control screen within 10 s (7.4% of trials), 2312 trials remained for analyses.

Visual search time and accuracy. Subjects were asked to respond as quickly and as accurately as possible. Accuracy was high (99%) and did not vary across blocks or levels of quadrant salience. As the salient quadrant was irrelevant to the goal of the task (i.e., uninformative of target location), there is no reason for it to influence search time. However, it is possible that crowding within the salient quadrant made search less efficient. To test the effects of quadrant crowding, we compared trial times from control trials to trial times when a salient quadrant (SQ) was present for those trials in which the first saccade was towards the quadrant that contained the target. To provide a strong test of the crowding hypothesis we compared control trials with strong saliency trials using Student’s paired t-test. This comparison showed no effect of strong quadrant saliency ($M = 1039.2$, $SE = 59.96$) on response times relative to control trials ($M = 1064.28$, $SE = 66.44$), $t = -0.841$; $p = .42$. This result effectively demonstrates that searching the salient quadrants did not affect trial response times relative to searching the control quadrants; hence, we found no evidence for potential confounding effects of quadrant crowding on response times.

\footnote{We thank an anonymous reviewer for pointing this out and suggesting this comparison.}
Effect of the salient quadrant. Two analyses examined whether initial saccades selectively favoured the salient quadrant. The first compared the proportion of initial saccades to what would have been expected by chance, and the second compares the proportion of initial saccades to the salient quadrant for the moderate and strong salience conditions. On average, 33.27% of initial saccades occurred to salient quadrants when one was present.

Two \( t \)-tests were used to compare the proportion of initial saccades to strong and moderate saliency quadrants to what would have been expected by chance (i.e., 25%). (All reported \( t \)-tests were two-tailed.) Initial fixations to strong saliency quadrants were significantly greater than chance levels (\( M_{\text{strong}} = 35.79\%, \ SE = 3.1 \), \( t(12) = 3.41, p < .01 \); however, moderate saliency quadrants were not (\( M_{\text{moderate}} = 27.58\%, \ SE = 2.5 \), \( t(12) = 0.920, p > .3 \). These results demonstrate a salient quadrant effect (pro-SQ) on initial saccades to strong saliency quadrants when the salient quadrant is task irrelevant.

Next, a \( 2 \times 4 \) repeated-measures ANOVA of salient quadrant level (moderate and strong) and block (1–4) was conducted on the proportion of initial fixations on salient quadrants. The ANOVA revealed a main effect of quadrant salience level, \( F(1, 3) = 8.295, p = .01 \), indicating that a greater proportion of initial saccades were directed to strong saliency quadrants than to moderate ones. Neither the main effect of block, nor the block \( \times \) quadrant saliency interaction, reached significance.

Saccadic reaction time. Endogenously influenced saccades have been demonstrated to result in longer saccadic reaction time (SRT) than exogenously controlled ones (Everling & Fischer, 1998; Godijn & Theeuwes, 2002; Hallett & Adams, 1980; Mulckhuys & Theeuwes, 2010; Nieuwenhuis et al., 2000; Wolfe et al., 2000). An important difference between the previous and current research is the potential for strategic variation within the salient quadrant paradigm. It is likely that participants are attempting to impose some endogenous strategy on the task to respond faster. Any “pattern-based” strategy (e.g., searching clockwise or counterclockwise from the top-left quadrant), history strategy (beginning a search on trial \( n \) within the same quadrant that the target was found on trial \( n - 1 \)), or any other strategy would be equally ineffective for finding the target and would lead to levels of initial saccades to the salient quadrant that are approximately equivalent to chance (i.e., 25%). Note as well, these examples of pattern-based strategies would be equally likely to start by searching a salient quadrant 25% of the time when one was present. Of course, it is also likely that on some trials participants adopt the strategy to begin search in the salient quadrant. For all of these reasons, some proportion of initial saccades to the salient quadrant will be endogenously influenced. Hence, although some pro-SQ saccades will be exogenously influenced, because of the
mixture of influences we did not expect strong differences in SRT between pro-SQ initial saccades and those to non-salient quadrants (non-SQ).

For trials containing a salient quadrant, we calculated SRT for pro-SQ vs. non-SQ initial saccades using the Vincentizing procedure (Ratcliff, 1979) to produce deciles for each participant, where the first decile contained the fastest SRTs and the tenth contained the slowest SRTs (see Figure 2). A $2 \times 10$ repeated-measures ANOVA on initial saccades (pro-SQ vs. non-SQ) and SRT decile was computed. As hypothesized, SRT did not differ as a function of initial saccade (i.e., pro-SQ vs. non-SQ), $F(1, 12) = 0.172, p > .60$. The initial saccade $\times$ SRT decile interaction was not significant, $F(9, 108) = 2.04, p = .07$.

The Vincentized SRT analysis is useful for uncovering main effects of salient quadrants on SRT, but does not provide any insight into the effect of experience on SRTs. To determine if SRT changed across blocks, and initial saccade type (pro-SQ or non-SQ), a $2 (\text{pro-SQ, non-SQ}) \times 4 (\text{block})$ repeated-measures ANOVA was conducted on SRT. The main effect of salient quadrant (pro-SQ vs. non-SQ) was not significant, $F(1, 12) = 0.22, p = .65$, nor was the main effect of block, $F(3, 36) = 1.62, p = .20$. The quadrant-salience effect $\times$ block interaction on SRT was not significant, $F(3, 36) = 1.64, p = .20$.

As mentioned earlier, it is likely that participants are attempting different search strategies to reduce response time and, as there is no reason to avoid it, all of these endogenous strategies would initially search the salient quadrant

![Figure 2. Experiment 1 saccade reaction time to salient quadrant and control quadrants by decile (Vincentizing procedure).](image-url)
25% of the time. In contrast, the finding of faster response times for exogenously influenced saccades is usually found in those situations where endogenous strategies work contrary to the exogenous influence (for example, see Godijn & Theeuwes, 2002; Mulckhuyse & Theeuwes, 2010; Nieuwenhuis et al., 2000). A test of this explanation will be presented in Experiment 2. In that study, subjects learn that search is faster if the salient quadrant is avoided and limit search to the remaining three quadrants. In the conditions of Experiment 2, as in the antisaccade task, most initial saccades to the salient quadrant should be exogenously controlled and SRT differences between pro-SQ and non-SQ initial saccades should occur.

Search costs of Experiment 1

The salient quadrant was irrelevant to the task; that is, it was not informative to the location of the target. However, response times associated with initial saccades towards and away from the salient quadrant when the target was within and outside of the salient quadrant can be computed to supply parameter values for our ideal observer model. When the target was in the salient quadrant (TQ = SQ), an initial saccade to the salient quadrant (pro-SQ) saved 564 ms of search time compared to when an initial saccade was to a non-salient quadrant (non-SQ). Search time differed significantly, $p < .0001$ (1110 ms pro-SQ vs. 1674 ms non-SQ), $F(1, 12) = 51.9$.

When the target was not in the salient quadrant (TQ $\neq$ SQ), an initial pro-SQ saccade added 239 ms to search time compared to when an initial saccade was to a non-salient quadrant. This difference in search time was stable and significant (pro-SQ = 1649 ms vs. non-SQ = 1410 ms), $F(1, 12) = 24.74, p < .001$.

Discussion

The results from Experiment 1 support our argument that the salient quadrant acts as an exogenous influence during search. First, if an endogenous strategy of initially searching the salient quadrant were imposed then equal numbers of pro-SQ initial saccades to the moderate saliency and strong saliency quadrants would be expected. (Indeed, this expectation is supported in Experiment 2.) In contrast, we find more initial saccades to the strong saliency than to the moderate saliency quadrant. Although it would be possible to amend the endogenous pro-SQ strategy hypothesis to treat moderate saliency versus strong saliency conditions differently, this would require additional (ad hoc) assumptions. In the absence of evidence to the contrary it strikes us as parsimonious to suppose that subjects attempted to apply endogenous strategies to all three conditions: control, moderate, and strong. That the distribution of initial saccades differs depending on the
degree of exogenous influence is strong support for the exogenous influence of the salient quadrant.

Second, if pro-SQ saccades were an effect of adopting an endogenous pro-SQ strategy we would expect the pro-SQ effect to vary with experience. Any endogenous strategy, other than one that explicitly examined the salient quadrant first, would be expected to show a pro-SQ effect 25% of the time. Consequently, as subjects try and reject successive endogenous strategies we would expect a minimum 25% of pro-SQ initial saccades. However, we documented a pro-SQ effect on 33.27% of trials containing a salient quadrant. We found no difference in SRT between pro-SQ and non-SQ saccades. Although we did not predict this result we do not find it surprising as past research showing faster response times to exogenous stimulus used paradigms in which the endogenous strategies that their subjects adopted conflicted with the draw of salient items. Under those conditions, it is easy to distinguish between exogenously and endogenously controlled saccades and test for SRT differences.

Based on the nature of the Experiment 1 task environment and the different degrees of salience, SRT may not be an adequate litmus test for detecting exogenously influenced saccades. We argue that other measures should be considered for determining the presence of exogenously influenced saccades, such as the likelihood of the effect relative to chance and objective measures of visual salience. Consequently, one should be cautious about concluding the presence or absence of exogenously influenced saccades within environments that do not restrict strategic variation and for which the exogenous influence is task neutral. However, if the salient quadrant becomes something to be avoided (i.e., task relevant), such as the abrupt onset in the antisaccade and oculomotor capture tasks, then we predict that subjects will eventually demonstrate SRT differences between pro-SQ and non-SQ saccades.

When initial saccades favoured a salient quadrant that contained the target, response times were reliably reduced by 564 ms. Based on the soft constraints hypothesis, if the salient quadrant was task relevant (i.e., reliably signalled the target location), then this half-second decrease in search time would have led subjects to discover and consistently apply a pro-SQ strategy. Conversely, an initial pro-SQ saccade added a 239 ms visual search cost when the salient quadrant did not contain the target when compared to an initial non-SQ saccade. If the salient quadrant were made task relevant by never containing the target (thereby effectively reducing the functional search space from four to three quadrants) then we would expect subjects to adopt strategies that avoid searching the salient quadrant. However, in speeded response tasks it is clear that deliberate strategies cannot always override exogenous influences (Everling & Fischer, 1998; Mitchell et al., 2002). Hence,
we would expect some oculomotor capture. This general prediction is the focus of our ideal performer model and is tested in Experiment 2.

An ideal performer model of exogenous influences on search efficiency

This subsection covers the development of a model of ideal human performance that can be used to predict human behaviour when changes are made to the salient quadrant paradigm. The ideal performer model is quite simple in form. The mathematical objective for the model is to maximize the efficiency of visual search; that is, find the target as quickly as possible, by optimizing the location and sequence of saccades. This objective of optimizing interactive behaviour through the minimization of time costs has previously been proposed as the soft constraints hypothesis (Gray et al., 2006), which states that over the 1/3 to 3 s time span, the mixture of perceptual-motor and cognitive resources allocated for interactive behaviour is adjusted based on cost–benefit tradeoffs, with time defined as the primary unit of cost. Within our analysis, trial times are considered costs, and are used to determine the utility of favouring or avoiding salient quadrants. Hence, the actions with the lowest costs (fastest expected trial results) are more often selected.

Previous researchers have also identified search efficiency as a useful performance objective for controlling eye movements (Najemnik & Geisler, 2005, 2008). In this previous work, the optimal locations for eye movements were determined based on template matching between a known search target and sensory information available in the periphery. By contrast, in our ideal performer model, eye movements are informed by statistical regularities in the task structure that can be used to predict the absence or presence of a search target at particular locations. Thus, our analysis suggests that not only are saccades influenced by uncertain information about a search target in the periphery (Najemnik & Geisler, 2005), but also that statistical information about where the target is not located also influences saccades.

In addition to the performance objective of minimizing time costs, our ideal performer model specifies two known constraints on human performance. First, it is assumed that visually salient locations exert an exogenous influence on oculomotor behaviour. That this is the case is a well-supported finding in the visual search literature (Lamy & Egeth, 2003; Wolfe, 1998; Yantis & Jonides, 1990), as well as the results from our first experiment. It is also the case that in most natural environments, fast and automatic saccades to visually salient locations are of biological utility—such a mechanism can facilitate detecting food items and warning signals from other organisms. A virtue of our experimental paradigm is that it enables quantifying the magnitude of this exogenous influence. In our first experiment, subjects
demonstrated a consistent bias towards the salient quadrant on their first saccade of each trial, even though it did not contribute to performance in the task. From the human behaviour we quantify the exogenous influence which is used as a constraint within the ideal performer model.

The second constraint on the ideal performer model is that on the 1/3 to 3 s timescale, interactive routines are selected in stochastic rather than deterministic fashion. This limitation on performance also has its origins in the soft constraints hypothesis (Gray et al., 2006). If actions are selected that minimize time cost, then a stochastic action selection mechanism will occasionally select actions that have a higher expected cost (lower utility). On the surface, it would seem that occasionally choosing the costlier of two alternatives would be a maladaptive strategy. However, this need not be the case (Müller & Krummenacher, 2006). Consider two strategies that differ in their expected cost by 100 ms—deterministically minimizing expected costs would demand choosing this slightly faster strategy 100% of the time. Yet this deterministic strategy may not be optimal for several reasons. First, it is known that the human ability to estimate time intervals is characterized by uncertainty (Gibbon, 1977). Thus, even if one action is reliably faster than an alternative, this difference may not be perceptible to the actor. Second, it neglects the fact that all organisms must learn the expected costs of actions through self-guided experience in inherently stochastic environments. Therefore the only means of acquiring accurate estimates of the expected costs associated with different actions is by occasionally sampling actions that are suboptimal according to current estimates. In the machine learning community, this tradeoff is known as the exploration-exploitation dilemma (Thrun, 1992), and in certain restrictive environments (Streeter & Smith, 2006), the optimal decision strategy can be derived analytically and shown to require sampling actions with higher expected costs.

With a mathematically well-defined performance objective in place, and the two performance constraints defined above (the exogenous influence of salience on visual search, and stochastic action selection), it is possible to derive boundedly optimal ideal performance for our visual search paradigm. In the following sections we derive our model and calibrate it to the behavioural data collected in Experiment 1 when the salient quadrant was task irrelevant. In the following section we use the calibrated model to predict a boundedly optimal level of oculomotor capture when the salient quadrant never contains the target, which will be the case in Experiment 2.

**Derivation of the ideal performer model**

The first step in formulating an optimal decision problem is to assign numerical values (or utilities) to the possible actions given the possible states of the world. In our simple analysis, states of the world correspond to whether or not the target is located in the salient quadrant. Thus, we can define a
binary indicator variable $t$ that takes the value 1 when the target is located in the salient quadrant, and 0 otherwise. For the space of possible actions, we consider two possible visual search routines: pro-SQ and non-SQ. A pro-SQ routine searches the salient quadrant with the first fixation. If the target is not located, the remaining three quadrants are searched in arbitrary order. The non-SQ routine does the opposite: the non-salient quadrants are searched first, and the salient quadrant is searched only if the target has not been located in one of the three nonsalient quadrants. The search routine chosen on a particular trial is indicated by the variable $a$, which is equal to either pro-SQ, or non-SQ. With two possible visual search routines and two possible states of the world, we construct a $2 \times 2$ utility matrix that assigns numerical values to each possible combination of world state and search routine.

To achieve this, we first need to define two relevant quantities: $\phi(n)$, which gives the time cost (in ms) associated with finding the target after searching $n$ quadrants, and $p(n|a,t)$, which gives the probability of having to search $n$ quadrants in order to find the target for a given world state $t$ and search routine $a$. With these two quantities defined, the expected search cost $U(a|t)$ of adopting action $a$ in a given world state $t$ is given by

$$U(a|t) = \sum_{n=1}^{4} \phi(n) \cdot (n|a,t)$$

This equation specifies the expected search cost, in terms of time, for a given world state and search routine. We derive the probability of finding the target after searching exactly $n$ quadrants, $p(n|a,t)$, as follows. The probability of finding a target after searching exactly $n$ quadrants is given by the probability of failing to find the target in each of the first $n-1$ quadrants searched, times the probability of successfully finding the target on the $n$th quadrant. For example, if the target happens to be located in one of the three nonsalient quadrants ($t=0$), and the searcher adopts a non-SQ search routine, then the probability of having a search length equal to one, $p(n=1|a=\text{non-SQ}, t=0)$, is simply the probability of correctly guessing the target location on the first try, which in this case is 1/3. Similarly, the probability of achieving a search length of two quadrants, or $p(n=2|a=\text{non-SQ}, t=0)$, is given by the probability of guessing wrong on the first try, and correctly on the second, or $(2/3) \times (1/3)$. Similar reasoning is used to define the probabilities for each search length for the other combinations of world state and search strategy. The complete probability distribution $p(n|a,t)$ is plotted in Figure 3(a).

The cost function $\phi(n)$ defines the search cost (measured in ms) associated with finding the target after searching $n$ quadrants. We derived
this function directly from our Experiment 1 data, and a linear model was fit to the empirical search times when the target was found after searching 1–4 quadrants (for this analysis, trials were classified by determining how many quadrants had been fixated at least once on that trial). The results of the model fit are shown in Figure 3(b). As can be seen, the empirical search times are accurately described by a simple linear function of the number of quadrants searched. These search costs were combined with the probabilities defined above to produce the table of expected search costs, $U(a|t)$. These computed costs are reported in Table 1.

For the ideal performer model, the true state of the world (whether the target is in the salient quadrant) is an unknown quantity on each trial. However, if the model possesses knowledge of the statistical relationship between the salient quadrant and the target location, it can use this information to minimize the expected search costs. This is incorporated into the model through a probability distribution over the world state, given by $p(t)$. For our first experiment, the target is located in the salient quadrant on 25% of trials, and so $p(t = 1) = 0.25$. This distribution is used to marginalize over the unknown world state, yielding the expected utility for each visual search routine, $U(a) = \sum_{t=\{0,1\}} U(a|t) \cdot P(t)$.

Figure 3. (a) The probability of having a search length of exactly $n$ quadrants as a function of visual search routine, $a$, and target location, $t$. When the salient quadrant does not contain the target ($t = 0$; top row) and the pro-SQ strategy is used (left column), then the probability of finding the target with the first saccade is zero (as the salient quadrant will be checked first). The probability of having a search length of $n = 2$ quadrants is given by the probability of failing to find the target on the first saccade (probability = 1), multiplied by the probability of finding the target on the second saccade (probability = 1/3). Similar reasoning is used to compute each remaining value. The bottom row shows the distributions over search length when the salient quadrant contains the target ($t = 1$), and the right column shows the distributions when a non-SQ search routine is adopted. (b) The search cost (measured in milliseconds) associated with finding the target after searching $n$ quadrants. Data is taken from Experiment 1. For the linear model, $r^2 = .9945$. 
With the expected utility of pro-SQ and non-SQ strategies defined in terms of search times, the only detail that remains is the mechanism for selecting between actions. Recall that one of the constraints imposed on the ideal performer is that on the 1/3 to 3 s timescale, a stochastic, rather than deterministic, process determines behaviour. We implemented this in our model using the soft-max equation, which is widely used as a mechanism for action selection in machine learning (Sutton & Barto, 1998; Thrun, 1992) as well as models of human choice (Anderson & Lebiere, 1998; Fu & Anderson, 2006; Gray et al., 2006):

\[
p(a) = \frac{\exp[-U(a)/\tau]}{\sum_{a_j \in A} \exp[-U(a_j)/\tau]} \quad (2)
\]

This equation gives the probability of selecting action \(a\), from among a set of possible actions \(A = \{\text{pro-SQ}, \text{non-SQ, non-SQ, non-SQ}\}\).\(^2\) The negative signs appear since utility gives a search cost that is to be minimized, rather than maximized. The parameter \(\tau\) controls the stochasticity in action selection: larger values of \(\tau\) increase the probability of selecting an action with higher cost. Although \(\tau\) is sometimes referred to as a “noise” parameter, as noted earlier, in any complex decision environment, occasionally sampling actions with higher expected costs is necessary for achieving optimal performance.

Rather than fitting the parameter \(\tau\) to the data from this experiment, we relied on the parameter estimate determined from a previous model of action selection in routine interactive behaviour (Gray et al., 2006). In this previous research, the task that was modelled was not visual search, but rather the selection of alternative strategies for encoding information in declarative memory. However, the parameter value used in that research was intended to reflect the properties of action selection in low-level interactive behaviour, and not just behaviour in the specific task environment. By reusing the same

\(^2\)Note there are four competing actions, one pro-SQ strategy, and three instances of non-SQ, corresponding to each of the three nonsalient quadrants.

<table>
<thead>
<tr>
<th>Target state</th>
<th>Search routine</th>
<th>Expected search cost (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t = 0)</td>
<td>(a = \text{anti-SQ})</td>
<td>1294.71</td>
</tr>
<tr>
<td></td>
<td>(a = \text{pro-SQ})</td>
<td>1631.92</td>
</tr>
<tr>
<td>(t = 1)</td>
<td>(a = \text{anti-SQ})</td>
<td>1969.13</td>
</tr>
<tr>
<td></td>
<td>(a = \text{pro-SQ})</td>
<td>957.50</td>
</tr>
</tbody>
</table>
parameter value it is possible to test the generality of this hypothesis with regards to its ability to account for human interactive behaviour on the 1/3 to 3 s timescale.

The final constraint placed on our ideal performer analysis was the exogenous influence of visual salience on the selection of fixation locations. We simply modelled this as an additive boost for the utility of the pro-SQ action. Thus, the utility of a pro-SQ search routine is given by $U(\text{pro-SQ}) + c$, where $c$ is a constant. We chose $c$ based on the data from Experiment 1. Recall that by the end of the first experiment, human subjects fixated the salient quadrant with their first saccade on 33.27% of trials (rather than the expected 25%), even though the salient quadrant offered no advantage in terms of locating the search target. We analytically determined the value $c$ that resulted in this empirically observed bias towards the salient quadrant. This value turned out to be $c = 197.68$. If costs are measured in milliseconds, then this constant is also defined in milliseconds, and can be interpreted as an exogenous influence of salience on search behaviour that acts as if an initial saccade to the salient quadrant somehow shaves 200 ms from the expected search time.

With all the parameters derived from Experiment 1 and previous research now in place, the ideal observer model produces a prediction that pro-SQ initial saccades will occur at a proportion of 33.27% when the target quadrant and salient quadrant are independent as observed in Experiment 1. This is not surprising given that the model was derived from Experiment 1 data; however, as discussed later, what we believe is surprising is the success of this model in predicting the results of Experiment 2.

**Ideal performer model predictions for task-relevant salient information**

In the first experiment, the salient quadrant was made task irrelevant by completely dissociating its location from the target location. Our ideal performer model provided a theory-based description of the results of Experiment 1. Our theory maintains that oculomotor capture reflects a boundedly optimal balance among the strength of exogenous influences, temporal cost of initial saccades, the cost of searching in four quadrants versus the savings from searching fewer, as well as the stochasticity inherent to the action selection mechanism. We now turn our theory-based model from description to prediction by using it to predict the probability of oculomotor capture when salient information can be used to infer where the target is not located.

The only change to our model is the probability that the salient quadrant contains the target. Whereas in our first experiment, $p(t = 1) = 0.25$, in our second experiment the salient quadrant never contains the target, and so
This change does not reflect a modification of the predictive model (Equation 2), but instead is used as a description of the task environment and reflects differences between experiments. As such, with this small change Equation 2 can be used to make predictions of oculomotor capture when the salient quadrant never contains the target. Intuitively, a strategy resulting in optimal performance would be one that always avoids the salient quadrant. However, the model predicts that the boundedly optimal probability of searching the salient quadrant is .20.

The model balances several simple factors. First and foremost, searching a quadrant that never contains the target adds a cost to search time. Second, the exogenous influence of the salient quadrant functions as if to decrease the costs of a search that includes the salient quadrant by 200 ms. Third, the determinism of temporal costs is modulated by noise in the cognitive, perceptual, and motor system. Without noise, the probability of searching the salient quadrant would be zero. The addition of noise raises this to .20. For our cost-structure approach, Equation 2 weighs the effect of the salient quadrant against the costs of a wasted search of that quadrant. Consequently, the model predicts that subjects will reduce the probability of searching a salient quadrant from .33 in Experiment 1 to .20 when the salient quadrant never contains the target, representing an attempt to avoid the salient quadrant (i.e., anti-SQ3 behaviour). Moreover, the model predictions are distinguishable from merely ignoring the salient quadrant, as just about any search strategy (clockwise, counterclockwise, zigzag, etc.) within the task environment would result in pro-SQ initial saccades at chance levels (.25).

The argument is that .20 is the ideal performance achievable by an integrated cognitive system whose performance is bounded by exogenous influences and system noise.

Model incorporating learning on the basis of experience

The ideal performer model previously described contained the unjustified assumption that subjects knew with certainty that the target would never appear within the salient quadrant. However, human subjects in our experiments were not instructed on relationships between the target location and the salient quadrant. Consequently, if humans used the salient quadrant as a structural cue to the location of the target, then they must have learned this relationship on the basis of their experience in the task. For example, if a subject has completed 20 trials of the experiment and never found the target in the salient quadrant, does this mean that the target will never appear in

\[ p(t = 1) = 0. \]

3Note our switch in terminology from non-SQ to anti-SQ. When initial saccades are under the control of exogenous forces we refer to saccades to nonsalient quadrants as non-SQ. In contrast, when an adopted strategy or statistical adaptation dictates the salient quadrants be avoided, we refer to this as an anti-SQ effect, similar to the antisaccade effect.
the salient quadrant, or is it just very unlikely? In our ideal performer model,
this is equivalent to uncertainty regarding the probability $p(t = 1)$. In this
section we extend our ideal performer model to rationally infer the relation-
ship between the salient quadrant and target location.

The optimal response to dealing with any form of uncertainty is to
maintain a probability distribution over the uncertainty, and update beliefs
using Bayesian inference. We therefore introduce the variable $\theta$, which
specifies the unknown probability that the target will appear in the salient
quadrant. From this, we have $p(t = 1|\theta) = \theta$, and we can define a prior
probability distribution over the unknown quantity, $p(\theta)$. After completing
each trial, the ideal performer observes the location of the target. This
information can be used by the model to update its beliefs about the unknown
parameter $\theta$, via Bayes’ rule:

$$ p(\theta|t) = \frac{p(t|\theta)p(\theta)}{\int_0^1 p(t|\theta)p(\theta)d\theta} \quad (3) $$

Equation 3 defines how the ideal performer’s belief regarding $\theta$ changes after
observing the outcome of each trial. For our analysis, we chose to model the
prior $p(\theta)$ using a Beta distribution. This was chosen mainly for mathematical
convenience, as in this case the posterior distribution, $p(\theta|t)$, also takes the
form of a beta distribution. Details regarding our choice of prior and the
derivation of the posterior distribution are provided in the Appendix.

For the learning version of the ideal performer model, the utility of each
action is determined by marginalizing over the unknown parameter $\theta$:

$$ U(a) = \sum_{t=\{0,1\}} U(a|t) \int_0^1 p(t|\theta)p(\theta)d\theta \quad (4) $$

Action selection is performed in similar fashion as the previous model,
using the soft-max equation (2). Learning curves for this model were
obtained by simulating two task environments: one that contained a chance
relationship between salient quadrant and target locations (i.e., a task-
irrelevant salient quadrant; Experiment 1) and one that contained a
perfectly negatively correlated relationship between salient quadrant and
target locations (e.g., a task-relevant salient quadrant) and running 1000
model subjects through each environment. A large number of model
subjects were simulated in order to obtain smooth learning curves. The
only learning component of the model was the updating of the belief
distribution regarding the location of the target at the end of each trial. The
model predicts no learning when the target and salient quadrant locations are independent, which was observed in the Experiment 1 human data. However, the model predicts that participants operating in an environment where the target is never located in the salient quadrant will begin with a pro-SQ effect similar to that observed in Experiment 1, but will eventually reach asymptote near an oculomotor capture likelihood of .20 (as established earlier).

**EXPERIMENT 2**

Experiment 2 tests the predictions of our model by disallowing the collocation of the target and salient quadrant (i.e., TQ ≠ SQ). As a consequence, the exogenously produced pro-SQ effect would lead subjects to initially saccade to a quadrant that is guaranteed to not contain the target. Unlike the first experiment where neither endogenous strategies nor statistical adaptations could perform better than chance, in the second experiment the target’s absence from the salient quadrant (i.e., TQ ≠ SQ relationship) could improve search efficiency by reducing the functional search space from four quadrants to three.

Our ideal performer model predicts that participants will begin with a pro-SQ effect equivalent to Experiment 1 and gradually reduce the effect to asymptote around .20 by the end of the experiment. The model balances the costs of searching the salient quadrant, exogenous influences, and inherent noise. However, unlike the model, in normal, interactive behaviour people are usually not aware of their eye movements (Hayhoe, 2000). Ballard has argued that processes below 1000 ms are typically beneath the threshold of conscious awareness; certainly, events occurring between 200 and 240 ms (the calculated savings from avoiding initial saccades to a salient quadrant) are well below the 1/3 s threshold for embodied cognition (Ballard et al., 1997). Hence, it is unclear whether this level of cost avoidance would suffice for participants to produce an anti-SQ strategy in accordance with the predictions of the ideal performer model.

Along with model predictions, we derived two other hypotheses from Experiment 1 results. First, if participants learn to avoid the salient quadrant, then differences between pro-SQ and anti-SQ SRT will emerge with experience as the anti-SQ saccades reflect the endogenous strategy of salient quadrant avoidance and pro-SQ saccades reflect the exogenous influence of the salient quadrant. Second, avoidance of the salient quadrant will not be perfect and will be modulated by the strength of the exogenous influence (i.e., better avoidance of moderate saliency than strong saliency quadrants).
Method

Paradigm. Experiment 2 used a $3 \times 4$ within-subjects design, with three factors of quadrant salience and four factors of block. The same sequence of events occurred on each trial as in Experiment 1 (see Figure 1). Subjects performed the same visual search task as subjects from Experiment 1; however, in marked contrast to Experiment 1, the target never occurred in the salient quadrant (i.e., $TQ \neq SQ$). As in the first experiment, subjects were told to minimize search time and maximize accuracy. They were told nothing about where to move their eyes. Importantly, subjects were not informed that the target was never located within the salient quadrant. As before, eye tracking was introduced as a gaze-contingent method for controlling gaze location at the beginning of each trial.

Subjects. Seventeen undergraduate students volunteered to participate. All subjects had normal or corrected-to-normal vision. One participant was later removed due to poor initial fixation control, leaving 16 subjects. The study lasted approximately 1 hour. Subjects were run individually and received course credit as compensation for their time.

Results

Results are divided between visual search time and accuracy, effect of the salient quadrant, and SRT analyses. Prior to conducting analyses, trials were excluded for which the gaze-contingent display did not result in an initial fixation on the fixation control screen within 10 s (8.9% of trials). These exclusions left 2787 trials for analyses.

Visual search time and accuracy. Subjects were asked to respond as quickly and as accurately as possible. By this measure, performance was high (99%) and did not vary across blocks or by conditions.

To determine the effect the new search environment had on subjects’ response times, a $3$ (salience level) $\times 4$ (block) repeated-measures ANOVA was computed. (Block violated the sphericity assumption and the Greenhouse-Geisser correction was used.) Only the main effect of block was significant, $F(1.36, 45) = 21.91, p < .01$; $M_{\text{block-1}} = 1822$ ms, $M_{\text{block-2}} = 1576$ ms, $M_{\text{block-3}} = 1385$ ms, $M_{\text{block-4}} = 1285$ ms, demonstrating that subjects improved their search times and were behaving quickly and accurately. No other effects were significant.

A saccade type (pro, anti) $\times$ block (1–4) repeated measures ANOVA was conducted on response times to determine if there was a response time benefit to avoiding salient quadrants. There was a main effect of saccade type, $F(1, 15) = 24.8, p < .001$, and a main effect of block (corrected for violating the sphericity assumption), $F(1.63, 24.5) = 21.75, p < .001$, indicating that trial
times were significantly greater when a pro-SQ initial saccade occurred than when an anti-SQ initial saccade occurred. These results demonstrate that SQ avoidance resulted in performance increases as measured by response time. No other effects reached significance.

**Effect of the salient quadrant.** To determine if subjects’ initial saccades were sensitive to the cost structure of the search environment, a 2 (salience level) × 4 (block) repeated-measures ANOVA was computed. A main effect of block was observed, $F(3, 15) = 6.32$, $p = .011$, where Block 1 had significantly more pro-SQ saccades ($M = 33\%$) than Block 4 ($M = 22\%$) supporting the hypothesis that subjects would learn to avoid salient quadrants, even though salient quadrants were demonstrated in Experiment 1 to induce a pro-SQ effect greater than chance. There was not a main effect of salient quadrant level, $F(1, 15) = 0.0002; p > .99$; nor was there a significant salient quadrant by block interaction, $F(3, 45) = 0.049, p > .98$.

A series of planned comparisons were conducted to determine if a pro-SQ effect was exhibited in Block 1 and if an anti-SQ effect was exhibited in Block 4. Collapsing over moderate and strong saliency levels, comparisons revealed a reliable pro-SQ effect in the first block, $t(15) = 3.01, p < .01; M = 33.5\%, SE = 0.02$; however, although a slight anti-SQ effect was observed by Block 4 ($M = 22\%, SE = 0.02$), it was not significantly below chance, $t(15) = -1.65, p > .10$.

Subjects initially fixated salient quadrants at a level above chance that was comparable to Experiment 1, and ended exhibiting a pro-SQ effect at levels below chance. These results support the predictions that under a TQ ≠ SQ paradigm, subjects would learn to avoid the salient quadrants with repeated exposure to the search environment.

Figure 4 shows the resulting learning curves for the ideal performer model compared to the empirical data. The human data has been aggregated into blocks of 24 trials in order to obtain smooth curves. The ideal performer model captures the main features of human learning (see Figure 4). In particular, the model is able to account for the decrease in pro-SQ fixations across the first few blocks of Experiment 2. Although the shape of the prior distribution $p(\theta)$ influences the slope of the learning curve, the model’s asymptotic predictions and main qualitative features are largely invariant to the precise choice of prior.

**Saccadic reaction time.** We calculated independent distributions of the SRT from initial saccades when a saccade was pro-SQ and anti-SQ (i.e., saccade type). Averages were produced using the Vincentizing procedure as in Experiments 1 and 2. A repeated-measures ANOVA with saccade type and decile as factors was computed. The latency differed between pro-SQ (276 ms) versus anti-SQ (304 ms) saccades, $F(1, 15) = 10.689, p < .01$. 

---

[The INSISTENCE OF VISION 1145](#)
The salient quadrant effect × decile interaction approached significance, $F(9, 135) = 2.446, p = .077$.

To determine if SRT changed as a function of experience with the environment (across blocks) and saccade type (pro-SQ or anti-SQ), a 2 (pro-SQ, anti-SQ) × 4 (block) repeated-measures ANOVA was conducted on SRT. As expected based on the Vincentized ANOVA, the main effect of salient quadrant effect (pro-SQ vs. anti-SQ) was significant, $F(1, 15) = 7.08, p = .018$.

More importantly, there was a significant saccade type × block interaction on SRT, where pro-SQ SRT decreased across blocks while anti-SQ SRT remained constant across blocks, $F(3, 45) = 3.64, p = .02$ (see Figure 6). As subjects gained experience in the task environment and reduced their proportion of endogenously controlled pro-SQ initial saccades, a differentiation in response time based on saccade type emerged. The lack of differentiation in Block 1 of Figure 6 is the same as was observed in Experiment 1; however, the differentiation in SRT by Block 4 is often used as a behavioural characteristic to distinguish between exogenously and endogenously influenced saccades (see Godijn & Kramer, 2006).

Discussion

Results from Experiment 2 supported our ideal performer model predictions and one of our hypotheses based on Experiment 1 results: (1) There was a
significant pro-SQ effect in the first block of trials in accord with the model, (2) the distribution of initial saccades changed with experience with the pro-SQ effect found in the first block vanishing into slight anti-SQ behaviour by the fourth block in accord with the model, and (3) initial saccade latencies varied as a function of the salient quadrant effect and task experience. As the number of initial fixations to salient quadrants did not vary with level of quadrant saliency, the results do not support the fourth hypothesis that the pro-SQ effect would be modulated by the strength of the exogenous influence imposed by salient quadrants.

Importantly, as subjects gained experience in the task environment, SRT became statistically distinguishable between pro-SQ and anti-SQ saccades by the last block (see Figure 6), resembling results from antisaccade (Everling & Fischer, 1998) and oculomotor capture tasks (Mulckhuyse & Theeuwes, 2010; Theeuwes et al., 1998). In common with Experiment 1, data from the first block suggest an exogenous influence of the salient quadrant. However, unlike that experiment, endogenous strategies that involve initial saccades to the salient quadrant are mitigated with experience of the relation-based scene statistics without instructions (Kramer, Hahn, Irwin, & Theeuwes, 2000) or cueing the location of the irrelevant, distracting information to be avoided (Munneke, van der Stigchel, & Theeuwes, 2008).

Avoiding salient quadrants reduced response times. The time it takes to saccade to a potentially informative quadrant is 434 ms if an initial pro-SQ is made versus 308 ms for an initial anti-SQ for a timesaving of 126 ms. Consequently, trial response times are faster when an initial saccade is anti-SQ ($M = 1460.05$ s) than when it is pro-SQ ($M = 1628.77$ s). This timesaving
was key to driving the boundedly optimal adaptations observed in Experiment 2.

The failure of our fourth hypothesis, which predicted greater oculomotor capture in strong versus moderate saliency quadrants suggests an interesting asymmetry between oculomotor capture, endogenous influences, and exogenous influences. In Experiment 1 the moderate saliency quadrants did not result in oculomotor capture at levels significantly greater than chance, yet in Experiment 2 participants were just as able to use the moderate saliency for avoidance as they were with the strong saliency quadrants. This underlying asymmetry must be further investigated to gain a better understanding between exogenous influences, endogenous influences, and oculomotor capture.

SUMMARY AND CONCLUSIONS

Our work seeks to predict search time by quantifying the exogenous attraction of salient information in terms of time. In Experiment 1, participants searched the salient quadrant first more than would be predicted by chance with the result being moderated by the level of visual salience. We termed this the pro-salient quadrant effect (pro-SQ). For Experiment 2, the salient quadrant never contained the search target. Subjects initially showed the same pro-SQ effect as in Experiment 1 but the effect was reduced with task experience. These experiments reveal two questions. First, why do people persist in making initial (or any) saccades to the salient quadrant when it never contains the target? Second, why saccade to the salient quadrant 22% of the time in Block
4, why not 5%, or 50%, or some other proportion? The ideal performer model provides an answer to each question.

The ideal performer model integrated the time costs of quadrant search, with a quantification of the exogenous influence of the salient quadrant, and an estimate of noise in the cognitive system. Based on Experiment 1 data, for Experiment 2 the model predicted a boundedly optimal solution of making initial saccades to salient quadrants on 20% of trials for the task investigated in Experiment 2. The prediction fell within confidence intervals from Block 4 of Experiment 2; hence, the answer to the two questions is that exogenous influences were mitigated by noise and experience, resulting in optimal performance given human processing constraints. Further, the mitigation was not immediate but took a substantial number of trials demonstrating that an endogenous anti-SQ strategy did not immediately take hold as demonstrated in Einhauser, Rutishauser, and Koch (2008), nor did the anti-SQ strategy ever take complete control as suggested by Foulsham and Underwood (2007).

The exogenous influence of the salient quadrant

There is ample evidence to conclude that the salient quadrant produces an exogenous influence on initial saccades. First, the greater attraction of the strong versus moderate salience quadrants supports the \textit{a priori} predictions based on Rosenholtz’s statistical saliency model (1999, 2001). Second, the pro-SQ effect did not vary with practice in Experiment 1 as might be expected of an endogenous effect but which would not be expected by an exogenous one.

Third, Experiment 2 made the salient quadrant task relevant to finding the target—the target was never within the salient quadrant. Consequently, an anti-SQ effect occurred by the fourth block. Importantly, there was an interaction of Blocks (1–4) and initial saccade location (pro-SQ vs. non-SQ) on SRT. The proportion of pro-SQ saccades in block 1 of Experiment 2 replicated the results of Experiment 1 and demonstrates the exogenous influence of the salient quadrant during search. By the fourth block of Experiment 2, the probability of an initial saccade to the salient quadrant decreased to below chance levels, suggesting the operation of an anti-SQ strategy for salient quadrant avoidance.

Ideal performer model

The ideal performer model is simple in its basic form (see Equation 3), as it does little more than moderate costs by noise (Maloney, 2003; Maloney, Trommershauser, & Land, 2007; Trommershauser, Maloney, & Landy, 2003a, 2003b). To this basic form we added very simple assumptions to calibrate the exogenous influence of the salient quadrant on initial saccades.
in terms of time. Importantly, the ideal performer model is solving for time, not for minimum oculomotor capture; however, the model predicts differences in oculomotor capture as a function of costs within the task environment in accord with the soft constraints hypothesis (Gray & Fu, 2004; Gray et al., 2006). The excellent fit of the model to the Experiment 2 results is interpreted as support for the soft constraints assumption that interactive behaviour at the 1/3 to 3 s level of analysis can be regarded as minimizing the use of the functional resource of time. Finally, the results of our model suggest that humans’ initial saccades are boundedly optimized to the search environment, and support previous research that humans are ideal observers given the limits of the human visuocognitive system (Najemnik & Geisler, 2008).

Making the assumption that all three factors in the ideal performer model are fixed (search costs, exogenous influence, and noise), what would be the implications of the model if, in Experiment 2, we told our subjects to never search the salient quadrant? It is clear to us that an endogenous strategy could be adopted that would reduce the number of initial saccades to the salient quadrant to a value closer to zero. However, this strategy would have to be one that compensated for noise in one of two ways. First, if subjects reacted slowly to the stimulus onset they should be able to show more precision in their initial saccades. For touching, this prediction follows directly from Fitts’ law (Fitts, 1954). For saccade control this prediction accords with the greater latency for endogenously controlled movements found in the antisaccade and oculomotor capture tasks. Second, also from Fitts (Fitts & Posner, 1967) we know that human performance becomes more accurate and faster with practice across a wide variety of tasks. However, it is unclear whether, with equal amounts of practice, an endogenously instructed group would ever match the response times of a noninstructed group (i.e., as per Experiment 2).

The learning process modelled in the ideal performer model could be extended in a number of ways. In our derivation, we have modelled learning of the structural regularities in the search environment, but not the temporal search cost per quadrant. Although human subjects most likely came in to our study with a lifetime of experience in searching for targets in the visual field, it is unlikely that they had much experience with the peculiarities and costs of our particular paradigm. It is also possible that the amount of noise in action selection, or the magnitude of the exogenous influence on search behaviour, could be adaptive processes that change with task experience. It is possible to extend our model to incorporate these features, but doing so would increase the complexity of the model without offering any additional insight as to the central theory underlying our analysis. Our ideal performer model is consistent with the assumption that human performance in the salient quadrant paradigm reflects ideal performance given the constraints of exogenous influences, noise in the cognitive, perceptual, and motor system, and the statistical and cost structures of the search environment (Najemnik & Geisler, 2008).
The dissociation between attention and oculomotor capture

In natural scenes, salient stimuli are often investigated to ensure that they pose no threat to our survival or safety. In such situations, oculomotor and attentional capture should be highly correlated. It also seems reasonable, that in safe and familiar environments, the attractive power of exogenous stimuli can be overruled (Einhauser et al., 2008; Foulsham & Underwood, 2007); that is, proactive control of attention can trump reactive control (Braver, Gray, & Burgess, 2007). Hence, in situations in which the information contained in a salient (i.e., exogenous) stimulus is not required for task performance, the exogenous stimulus can be preattentively ignored, and neither oculomotor capture or attention capture need occur.

In contrast to the tasks investigated by Einhauser et al. (2008) and Foulsham and Underwood (2007), spatial cueing paradigms require either some attention to the cue or use of preattentive information about the cue, but not oculomotor capture. In antisaccade tasks, subjects are explicitly instructed to avoid the cue and to saccade to the opposite side of the screen, the successful performance of this task shows that oculomotor capture can be dissociated from attention or preattentive information. It seems that this dissociation requires some effort and is more difficult for some clinical populations (Everling & Fischer, 1998), the elderly (Butler, Zacks, & Henderson, 1999), and the habitually error-prone (Larson & Perry, 1999).

In contrast to the antisaccade task, for the salient quadrant task we did not explicitly instruct subjects to avoid the salient quadrant. As we claim that our ideal performer model illustrates a near-optimal adaptation to the cost structure (in time) of the task environment, we find ourselves required to make an interesting prediction. If we can quantify the exogenous influence of the cue in the antisaccade task in a paradigm where the cue is task irrelevant (i.e., uncorrelated with the target location), then we can derive an ideal performer model for the antisaccade task for those cases in which subjects are told to minimize performance time but are given no instructions with regard to their saccades (i.e., when they are not told to avoid the cue). Furthermore, we would expect some tradeoff between mean response times and rates of oculomotor capture.

Why quantify search time?

Quantifying the factors that influence visual search time is an intriguing basic research problem that may advance our understanding of cognitive control (Braver et al., 2007; Cooper, 2010) and serve to integrate the theories and findings of visual search into overarching architectures of cognition (Gray, 2008). It can also increase the profile of our efforts by relating basic research to the emerging importance of visual search for cognitive engineering (Gray, Schoelles, & Myers, 2003).
Increasingly, complex computational cognitive models are being applied to study how people locate information on a webpage. Much research focuses on the semantic information contained on the pages or in the links (Fu & Pirolli, 2007; Teo & John, 2010); however, none use measures of visual salience even though salience can dominate semantics in the first 1.5 s of search in an unfamiliar scene (Carmi & Itti, 2006). These same saliency factors are important in other complex visual searches conducted under time pressure such as those involving airport security (McCarley, Kramer, Wickens, Vidoni, & Boot, 2004; Wolfe et al., 2007). Likewise, the visual presentation of massive amounts of abstract information, visual analytics (Keim, Robertson, Thomas, & van Wijk, 2006; Lee et al., 2006; Thomas & Cook, 2005) has become a concern for many basic researchers from fields such as physics and chemistry as well as from national security agencies. Models predicting performance in these complex task environments must be informed by theories of the tradeoffs between exogenous and endogenous influences, and one way is to quantify those tradeoffs using the currency of time.

REFERENCES


APPENDIX: DERIVATION OF THE BAYESIAN LEARNING MODEL

For the learning version of the ideal performer model, the agents had to learn the statistical relationship between the salient quadrant and target location. In our model, this relationship is governed by the parameter $\theta$. In reality, during Experiment 1 there was a 25% chance that the salient quadrant would contain the target, and thus $\theta = 0.25$, whereas in Experiment 2 the target never appeared in the salient quadrant, so that $\theta = 0$. Subjects in the experiment, however, were not given any information about the relationship between the salient quadrant and target location. Faced with this uncertainty, the optimal response is to maintain a probability distribution over the unknown parameter, $p(\theta)$. For our model we assumed a prior distribution over $\theta$ given by a beta distribution with parameters $\alpha$ and $\beta$:

$$p(\theta) = \text{Beta}(\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1 - \theta)^{\beta-1}$$ (A1)

The beta distribution is commonly used as a prior over a probability as its support is restricted to the range $[0, 1]$, and also due to the fact that it is the conjugate prior to the binomial distribution, meaning that in our case the posterior distribution $p(\theta|t)$ also takes the form of a beta distribution. In particular, after observing the outcome $t$ of each trial, the posterior distribution over $\theta$ is given by

$$p(\theta|t) = \begin{cases} \text{Beta}(\alpha + 1, \beta), & t = 1 \\ \text{Beta}(\alpha, \beta + 1), & t = 0 \end{cases}$$ (A2)
The marginal distribution governing the probability that the target is located in the salient quadrant also has a particularly simple form,

\[ p(t = 1) = \frac{\alpha}{\alpha + \beta} \]  

(A3)

In specifying the initial parameters \( \alpha \) and \( \beta \), a number of choices are possible. A uniform distribution over \( \theta \) results when the prior distribution is initialized with \( \alpha = \beta = 1 \). This corresponds to the case where the ideal performer model assumes that all values of \( \theta \) are equally likely. Alternatively, one might assume a priori that the salient quadrant is more likely to contain the target than the other quadrants, or less likely. In implementing our model, we chose to bias the ideal performer towards believing that the target location was independent of the salient quadrant, or that \( \theta = 0.25 \) is the most likely value. We implemented this by choosing \( \alpha = 5, \beta = 15 \). Note that with this parameterization, the belief distribution still allows for the possibility that the target will always appear in the salient quadrant, or alternatively that the salient quadrant will never contain the target; this is due to the fact that the probability densities for \( p(0 - 0) \) and \( p(\theta = 1) \) are nonzero. Figure A1 compares the shape of the prior distribution for two cases. In developing and testing the ideal performer model, it was found that the key predictions from the model are largely invariant with respect to the precise choice of parameters for the prior distribution.

**Figure A1.** Comparison of two different parameterizations of the beta distribution. (A) Uniform distribution. (B) Prior distribution biased towards believing that the target location is independent of the salient quadrant.