

RICE UNIVERSITY

Visual Displays: Developing a Computational Model Explaining the Global Effect

by

Clayton Stanley

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE

Master of Arts

APPROVED, THESIS COMMITTEE:

Michael D. Byrne
Assistant Professor, Psychology

James R. Pomerantz
Professor, Psychology

James L. Dannemiller
Lynette S. Autrey Professor, Psychology

HOUSTON, TEXAS
MAY 2009

ABSTRACT

Visual Displays: Developing a Computational Model Explaining the Global Effect

by

Clayton Stanley

This work aims to integrate Byrne's theory of visual salience computation (2006) with Salvucci's model of eye movements (2001) by testing participants on a visual search task similar to Findlay (1997). By manipulating the number, salience, and spacing of targets, participants exhibited the global effect averaging phenomena during the first recorded saccade, whereby short-latency saccades land in between adjacent objects. Previous work has argued that the saccadic targeting system causing the averaging is influenced both by the salience and arrangement of objects displayed (Rao, Zelinsky, Hayho, & Ballard, 2002). However, to accurately account for these results, we did not have to couple the salience system with the saccadic targeting system. Instead, the systems work sequentially and in isolation, whereby the salience system simply hands off the next object to examine to the targeting system, whose accuracy depends only on saccadic latency and the location of the targeted and non-targeted items.

TABLE OF CONTENTS

LIST OF TABLES.....	4
INTRODUCTION.....	5
Saccade Latency vs. Accuracy: The Speed/Accuracy Tradeoff.....	11
EXPERIMENT.....	14
Method.....	16
Results.....	26
MODEL.....	48
Towards Modeling Effects.....	48
Model Results.....	68
DISCUSSION.....	80
REFERENCES.....	87

LIST OF TABLES

Table 1. Number of distractors sharing color (1 st term in parentheses) & shape (2 nd term in parentheses) for each experimental condition.	19
Table 2. Number of observations in each cell of the experimental design.	37
Table 3. Effect of experimental manipulations on saccade distance.	39
Table 4. Effect of experimental manipulations on saccade directional error.	39
Table 5. Effect of experimental manipulations on adjusted saccade latency.	50
Table 6. Free parameters used in the targeting system to model data.	56
Table 7. Free parameters used in the Byrne's salience equations to model data.	61
Table 8. Free parameters used in EMMA to model data.	63
Table 9. Model performance with and without trimming < 100 ms saccade latencies.	79

INTRODUCTION

Imagine someone looking for the keys to his car. It's morning, and he remembers leaving them on his desk the night before, so he starts by looking there. The car keys are attached to the keychain he uses everyday, so he can easily envision what the keychain looks like. The color of his desktop is charcoal, and most of the other items sitting on the desk are dark colors. However, one object with bright red coloring jumps out at him. He remembers the keychain has a bright red rectangular shopping card attached to it, so he looks over at the object. After getting a closer look, he recognizes that the object is indeed his keychain, grabs his keys, and moves on his way.

Visual search tasks like this are performed every day of our lives, and several coarse behaviors of the vision system during these tasks are apparent:

[1]: If one knows the target object is located in a subset of the visual field, one generally looks for the object in that subset. In the example, the man started looking for the car keys on his desk, the last place he remembered leaving them.

[2]: If one knows the target object has certain attributes associated with it (e.g., color, shape, texture), one generally looks towards objects with these attributes. The man had an idea what the keys looked like, and his glasses looked somewhat different, so he quickly glanced over his glasses and looked elsewhere on his desk for his keys.

[3]: One generally looks towards objects in the visual field that are less like the others (e.g., a red rectangle in a sea of green rectangles is highly salient and will quickly attract attention). The bright red rectangular shopping card on the man's

keychain quickly attracted his attention since the rest of the desktop consisted of darker colors.

However, coarse behaviors are alone not sufficient to predict variables such as eye movements and search times that occur at the second and millisecond level. A more finely-grained model of the visual system must therefore be derived in order to predict these observations at a high level of accuracy. Nevertheless, a finely-grained model must still capture macroscopic observations when analyzed at a coarse level. Our goal therefore of building a model of the visual system is this: include enough components that visual search times and precise eye movements can be accurately predicted while ensuring that the model still allows the observed coarse behaviors of the visual system to emerge.

To model visual search behavior, we try to answer the following two questions as accurately as possible:

[1]: Given the salience of objects generated by both bottom-up and top-down cognitive processes, to where does one attend next?

[2]: How does where one currently attends affect the salience of objects displayed?

Answering these questions will allow one to model a typical visual search task as an iterative process where the individual loops through these questions until the target object is found.

Bottom-up cognitive processes. In a visual field containing numerous objects, many features (e.g., color, shape, size, texture) are processed early, in a parallel manner, and in an amount of time that is roughly invariant to the number of objects

in the field (Treisman & Gelade, 1980; Wolfe, 1994). This information is the same regardless of task demands and is commonly referred to as “bottom-up activation” in the visual search literature. That is, each object in the visual field receives a certain amount of salience (i.e., activation) for its features, regardless of task demands, that is processed for all objects in the visual field in a roughly parallel manner.

But how much salience should each object receive for its features? One could argue that the visual system is solving a resource allocation and limited bandwidth problem by being salience-sensitive. Resources are limited because humans have only one set of eyes and severe acuity limitations over most of the visual field. Acuity limitations would then drive the eyes to move around and sample from a probabilistic environment, attempting to maximize the amount of information passed through the system per unit time. To behave optimally, objects with the highest information content (i.e., objects unlike the others across one or many features) should then be looked at first. This allows one to argue that the information content of each object *is* each object’s activation or salience.

Using this approach, an object’s salience is determined strictly by how much information could be extracted if attentional resources were devoted to that object. The amount of information contained in an object can certainly be dependent on basic perceptual properties causing a response to the stimuli (e.g., if the object is really bright, it probably carries a lot of information). However, salience derived from an object’s information content acts further downstream than these basic perceptual properties, and aims to place higher salience values on more *distinctive*

objects (e.g., the one slightly dim object placed in a sea of bright objects), and not necessarily the most *intense* object.

Byrne (2006) developed a mathematical model using this hypothesis to calculate salience values. For bottom-up attentional processes, he uses the Hick-Hyman law to determine the amount of information each object contains. This bottom-up activation component aims to account for the 3rd coarse behavior previously discussed.

Various properties of the eye are not consistent throughout all viewing eccentricities. For example, the photoreceptor density varies across the retina (Hirsch & Curcio, 1989), and the retinal surface thickness across the eye is thinnest near the fovea (Polyak's study, as cited in Findlay & Gilchrist, 2003). This in turn sweeps away overlying neural layers and gives the fovea better acuity (Anstis, 1974). Because of these acuity differences between the fovea and periphery, how salient an object appears to someone may depend on the current place he/she is looking (i.e., the salience of objects may be fovea-dependent).

Top-down cognitive processes. Additionally, the task the individual is performing will affect what she chooses to look for in the scene. This task-dependent information, a function of high-level cognitive processes, is commonly referred to as "top-down activation" in the visual search literature. Byrne (2006) accounts for these top-down processes by increasing the information content of an object for each high-level task constraint it meets (e.g., the target object must be red, it has to be a rectangle, it has to be located in the top-right corner of the display). This allows the individual to have some cognitive control over what she wishes to

attend to. This top-down activation component aims to account for the 1st and 2nd coarse behaviors previously discussed.

Where does one attend to next? The salience of objects generated from both bottom-up and top-down cognitive processes, when aggregated, form a unique salience map corresponding to the amount of activation inherent in each location in the scene. Determining the next attending location is then done by competing neurons on the salience map. This process is stochastic, but on average returns the location corresponding to the highest amount of activation as the next location to attend to (Itti & Koch, 2000; Koch & Ullman, 1985; Logan, 1996; Pomplun, Reingold, & Shen, 2003; Rao et al., 2002; Wolfe, 1994).

Byrne (2006) formalized this process of combining bottom-up and top-down activation to form the salience map by simply adding all the partial activations together. Specifically, Byrne's salience model assumes that the salience L of object i in the visual field is defined as

$$L_i = \sum_{k=1}^{\#attr(i)} \log_2 \frac{1}{p_i(v_k)} \gamma_k + \log_2 \frac{1}{p_i(v_s)} + \sum w_j S_{ji} + \varepsilon \quad (1)$$

where the first term is the object's bottom-up salience, which is the sum of the information content contained in each attentional channel (e.g., color, size, shape), and is calculated for each object immediately after a scene is displayed (i.e., in a parallel manner). The second and third terms correspond to each object's top-down activation based on spatial and value guidance respectively. These values are calculated before each movement of visual attention as well, and depend on what the individual is currently looking for in the scene (e.g., a large red cross, probably

located in the top-right corner of the display). The last term captures the noise inherent in the system when an individual calculates these salience values. This is modeled using a logistic noise distribution with a mean of 0 and a configurable s -value. Using Equation 1, when a shift of attention is requested, attention will move towards the object with the current highest salience (i.e., towards the object with the highest L_i).

Attention affecting the salience of objects. Visual attention can drive eye movements, which is incorporated into Eye Movements and Movement of Attention (EMMA), which is implemented in ACT-R (Salvucci, 2001). The amount of time to encode an object displayed in the visual scene is dependent on two factors: how frequently the object has been encountered (Schilling, Rayner, & Chumbley, 1998) and the eccentricity between the object and the fovea (Rayner & Morrison, 1981). Specifically, the EMMA model assumes that the time T_{enc} to encode object i is defined as

$$T_{enc} = c \cdot [-\log f_i] \cdot e^{d_i} \quad (2)$$

where f_i is a measure of each object's frequency of occurrence in the external world, d_i the eccentricity distance between the object and the fovea, and c is a scaling constant.

When attention moves towards the most salient object, the object's attributes will start being encoded. Additionally, a saccade will be requested to move the fovea closer to the object, so that the object's attributes can be encoded at a faster rate (reducing d_i in Equation 2 reduces T_{enc}). When an eye movement is requested, time is required to both prepare and execute the saccade. The EMMA model assumes that

these stages are noisy (sampled from Gamma distributions) and take around 150 and 20 ms respectively. An additional amount of time is added to the execution stage depending on how far the saccade actually has to go (i.e., 2 ms for each degree of visual angle). If the object has not been completely encoded before the preparation stage completes, the eyes will move to help reduce T_{enc} . Otherwise, the saccade will be cancelled.

In this manner, attentional processes drive eye movements (i.e., the eyes follow visual attention, driven by the desire to maximize the rate of information processed at any given moment). Additionally, since the current location of the fovea affects the salience of the objects on the display, and visual attention drives the eyes to move, where one currently attends ultimately influences each object's salience.

Saccade Latency vs. Accuracy: The Speed/Accuracy Tradeoff

When a visual display contains multiple objects arranged relatively close to each other, saccades executed quickly (i.e., with latencies less than ~ 300 ms) often land in between two adjacent objects, while saccades with longer latencies are more accurate and generally fall on one object or the other (Findlay, 1997; Ottes, Gisbergen, & Eggermont, 1985). This effect (i.e., inaccurate saccades with short latencies) has been referred to as the “averaging phenomenon” (Ottes et al, 1985) or “global effect averaging phenomenon” (Findlay, 1997) in the literature.

Additionally, Rao et al. (2002) analyzed the first 3 saccades towards a target object and found that the first saccade moved approximately to the center of the display, while the second and third saccades spanned outward from the center

honing closer and closer to the targeted object's location. They referred to this effect as the "center of gravity phenomenon" (Rao et al., 2002).

The main difference in the task analyzed by Rao et al. (2002) compared to Findlay (1997) and Ottes et al. (1985) is that Rao's task had objects arranged at a much higher eccentricity from the initial fovea location, causing more than one saccade to be required to reach the targeted object. Nevertheless, both cases observed short-latency saccades falling in between multiple objects, and this effect merely started out on a larger scale for Rao's case because of the increased starting eccentricity. In a general sense, the commonality between the "center of gravity phenomenon" and "global effect averaging phenomenon" is that a speed/accuracy tradeoff exists between initiating a saccade quickly and having that saccade accurately land on a targeted object.

Rao et al. (2002) argued that this speed/accuracy tradeoff occurs because the salience map takes about 400 ms to evolve. Eye movements are then directed towards the vector average of objects on the display weighted by this salience map. As time moves forward, targeting accuracy improves because the signal to noise ratio of the target compared to the distractors increases, meaning that the target pulls harder than the rest of the objects on the display to move the eyes to its location. If modeled this way, Byrne's salience equation (i.e., Equation 1) becomes temporally dependent, where the starting salience of the target is equal to the distractors, and the signal to noise ratio increases as time moves forward. However, it is unclear how the EMMA model would fit (i.e., Equation 2) within this context,

since the model assumes that the visual system is trying to encode a particular object, and not the entirety of the scene all at the same time.

An alternative hypothesis is that the speed/accuracy tradeoff occurs because the targeting system itself takes a bit of time to program a saccade accurately towards the targeted object. In this way, the salience of objects is calculated almost instantaneously, has little to no temporal component, and acts prior to and separate from the targeting system (i.e., the salience system provides the input to the targeting system specifying which object to look at). With this hypothesis, the initial targeting activation for targeted and non-targeted objects is roughly equal, and as time moves forward, targeting accuracy improves because this signal to noise ratio of targeted to non-targeted objects increases, causing the targeted object to pull the saccade harder towards its location compared to the non-targeted objects. Salvucci (2001) suggests incorporating an undershoot bias into EMMA's targeting system for future revisions, presumably to account for these "center of gravity" style saccades. Modeling this hypothesis would simply replace EMMA's current targeting system with a temporally-dependent system, while leaving the system calculating the salience of objects (i.e., Equation 1) temporally independent.

To tease apart these two hypotheses, an experiment was designed where the signal to noise ratio of target to distractors was manipulated alongside the amount of time that objects had been displayed. If the salience map takes about 400 ms to evolve, having the objects (both target and distractors) displayed for approximately 1000 ms before searching for the target object should dramatically improve targeting accuracy (i.e., less eye movements in between adjacent targets).

Additionally, if the salience of the target is high, eye movements should land closer to the targeted object as well, since a reasonable signal to noise ratio between targeted and non-targeted objects would take less time to reach. Conversely, if targeting accuracy is roughly independent of the objects' age, one could argue that the speed/accuracy tradeoff is caused by a system programmed after the target to find has been encoded (e.g., in the saccade targeting system). Further, if eye movement accuracy towards the targeted object is roughly independent of the object's salience, one could argue that this targeting system isn't largely concerned about what the object is, and instead cares only about where the objects are and which object it's been told to look at.

EXPERIMENT

Findlay (1997) recorded eye movements and search times during a conjunctive search task where objects were placed on two concentric rings. Erroneous first saccades (i.e., initial eye movements towards an object other than the target) were more likely directed toward distractors sharing either shape or color with the target object. Subjects were therefore using task-dependent information (i.e., the shape and color of the target object) to guide their search.

However, the salience of the target relative to distractors was not manipulated. To tease apart the two hypotheses discussed, the signal to noise ratio must be altered alongside the objects' age. We therefore manipulated the target signal to noise ratio by changing the ratio of distractors sharing the target's color to distractors sharing the target's shape, allowing for the distractor ratio effect (e.g.,

Bacon & Egeth, 1997; Shen, Reingold, & Pomplun, 2003; Stanley & Byrne, 2008) to emerge.

Findlay (1997), as part of a separate experiment, used one concentric ring and manipulated the relative distance between two target objects (i.e., next to each other or one space apart). This design caused short-latency saccades to often fall in between adjacent objects. However, these results were collected separate from the conjunctive search task just described (i.e., two separate experiments). It would be beneficial to combine these manipulations into a single experimental design so that possible interactions could be analyzed. A useful question one may then be able to answer is: are eye movements towards the center of two adjacent objects more common if those objects have a higher relative salience compared to the other objects on the screen? We therefore manipulated the relative distance between two target objects alongside manipulating the ratio of distractors sharing color and shape.

As previously discussed, a bifurcation arises around 300 ms where afterwards eye movements are directed towards either of the two targets but not in between. If this phenomenon occurs solely because the salience map takes about 400 ms to evolve, after the objects have been displayed for 400 ms, saccades should be much more accurate at landing on targeted objects. To fully challenge this account, we used three different “presentations” of the task, aimed to manipulate the age of the objects displayed: the first displayed the target to find and then the concentric rings, the second displayed the concentric rings and then the target to find, and the third displayed the concentric rings first and then the target to find

whilst swapping the location of the objects on the rings. If the phenomenon occurs because the salience map is evolving, eye movements should be much more accurate for the second case and a speed/accuracy tradeoff should exist primarily for the first case. The third case was designed to test how well and how long participants remember where specific objects resided before the objects are moved.

All of these manipulations (i.e., ratio of same shape/color distractors, relative distance between two target objects, task “presentation”) were tested for each subject, shuffled randomly, and given during the course of a single experimental session. Both eye-movement and response-time data were recorded, and to account for the results, a fully integrated analysis of the attentional processes involved in the task was required.

Method

Participants

38 Rice University undergraduates (55.3% female) participated to fulfill experiment participation requirements for their psychology classes. The participants had a mean age of 19.8 years (3.0).

Apparatus

Stimuli for the experiment were displayed on a 19” Dell CRT monitor at a frame rate of 75 Hz. Subjects were placed directly in front of the display and interacted using standard keyboard movements. The same machine was used for all participants.

The eye tracker used was an ISCAN RK726/RK520 HighRes Pupil/CR tracker with a polhemus FASTRACK head tracker. Head-mounted optics and a sampling rate of 60 Hz were used in the experiment.

Stimuli and Design

The displays consisted of 19 items. Each item was 1 of 4 possible shapes (i.e., square, cross, circle, triangle) and 1 of 4 possible colors (i.e., the following CIE coordinates (measured with a Minolta meter) Black: $x=.41, y=.18, Y=1.1$; Red: $x=.49, y=.29, Y=14.2$; Green: $x=.35, y=.41, Y=28.8$; Blue: $x=.24, y=.13, Y=6.7$). All items had equal surface area. Each item (except for the target to find placed in the center of the display) was placed on either the inner or outer concentric ring, arranged radially, and spaced equally apart. The inner and outer rings had 8 and 10 items respectively. The display was viewed from a distance of approximately 60 cm, and the eccentricity of the inner and outer rings was 5.7 and 10.2 degrees. Finally, for each trial all objects on the display were randomly rotated by some amount (e.g., all crosses and squares for a trial may be oriented with tips at 0, 90, 180, 270 degrees or rotated by 11 degrees). An example of the display during the search phase of the experiment is included in Figure 1.

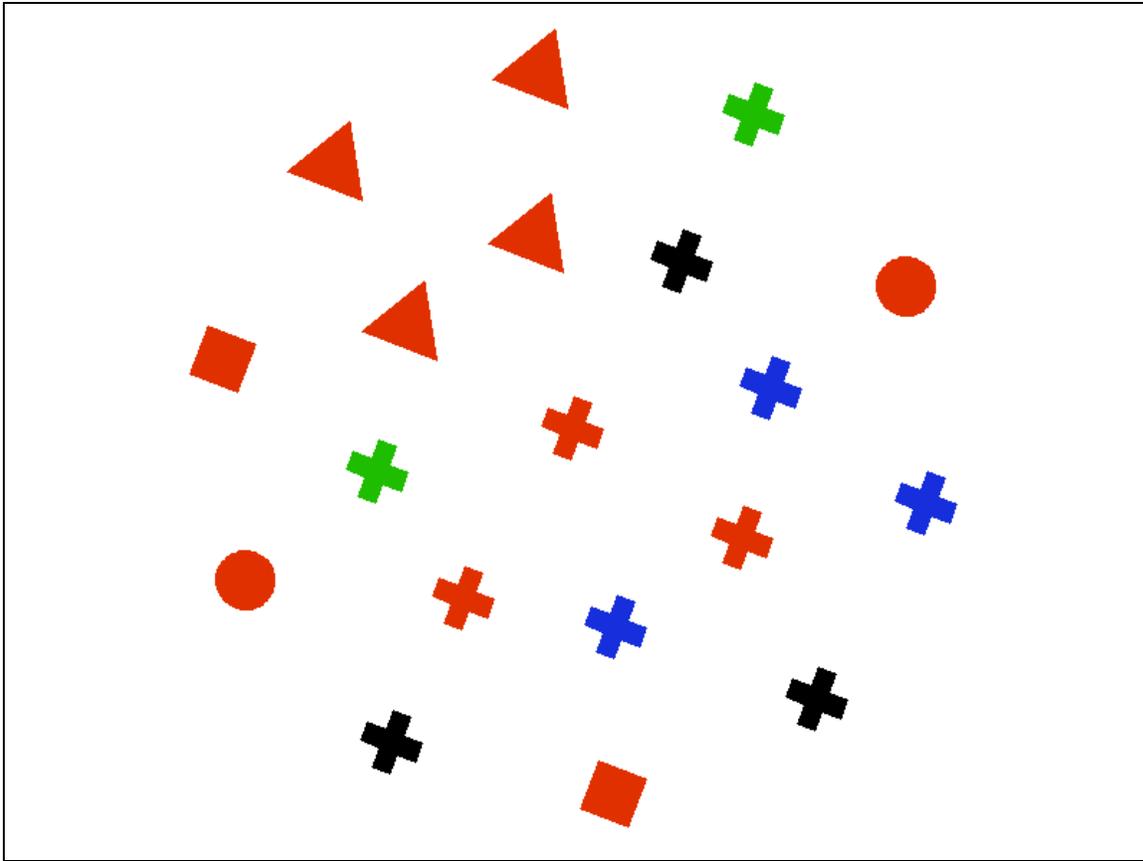


Figure 1. Example of the display during the search phase.

Five independent variables were manipulated: target present/absent, ratio of color/shape distractors, number of target objects, relative spacing between target objects (if two targets are displayed), and “presentation” of the task).

Target present/absent. For half of the trials, the target object was located on one of the two concentric rings. For the other half of the trials, the target object was absent from these rings.

Ratio of color/shape distractors. The distractor ratio effect causes quick response times at the fringes (i.e., low number of distractors sharing either color or shape) and slow response times in the middle (i.e., equal number of distractors sharing color and shape). Therefore, three data points were taken across this

dimension to capture this quadratic effect: a high ratio of color/shape distractors, an equal ratio of color/shape distractors, and a high ratio of shape/color distractors. Because we manipulated both the number of targets displayed and if the target was displayed, the exact ratio of color/shape distractors was slightly dependent on other manipulated dimensions. Specifically, the exact number of distractors used for each condition is included in Table 1.

Table 1. Number of Distractors Sharing Color (1st Term in Parentheses) & Shape (2nd Term in Parentheses) for Each Experimental Condition.

Condition	Less Sharing Color	Less Sharing Shape	Equal Sharing
2 Targets & Present	(3,13)	(13,3)	(8,8)
2 Targets & Absent	(3,15)	(15,3)	(9,9)
1 Target & Present	(3,14)	(14,3)	(9,8) or (8,9)
1 Target & Absent	(3,15)	(15,3)	(9,9)

Number of target objects. One or two targets were displayed on the screen.

Relative spacing between target objects. If two targets were displayed, four different conditions were sampled: targets located on the inner ring and placed side by side, targets on the outer ring and placed side by side, targets on the inner ring placed with one distractor in between, and targets on the outer ring with one distractor in between. This dimension aimed to test how the frequency of observed short-latency saccades depends on the spacing of target objects.

Task "presentation." A trial consisted of two phases: presenting the target to find and presenting the 18 objects on the concentric rings. However, the task was designed such that a participant could not deduce if the target object was present/absent by merely looking at the 18 objects on the concentric rings. He needed to know which target he was looking for before determining if that target

was present in the scene. This allowed us to manipulate the order of the two phases without affecting response times due to high-level cognitive processes. Data was therefore collected for both phase orderings (i.e., present the target to find and then the concentric rings, present the concentric rings and then the target to find), as well as an additional condition where the objects on the concentric rings were presented first and then the target to find whilst swapping the location of the objects on the rings. For this condition, target object(s) kept their relative spacing and remained on the same ring after shuffling. However, distractors were allowed to swap rings and alter their relative location to the target object(s). The duration of the first phase presented was sampled from a continuous uniform distribution ranging between 500 and 1500 ms.

These manipulations produced the following factorial design: 2 [target present/absent] x 3 [distractor ratio] x (1x1 + 1x4) [one or two targets & relative spacing between targets if two targets] x 3 [task presentation] = 90 cells. 4 repetitions for each cell were taken. Therefore, a total of 360 trials were shuffled and presented to each participant. The duration of the experiment was approximately 1 hour.

Procedure

Participants were asked to determine if a target object was present/absent from a visual scene. They placed each hand on the home row of the keyboard and indicated if the target was present/absent by pressing “j”/“f” respectively.

If the target to find was presented before the concentric rings, this object was displayed in the center of the screen, and after between 500 and 1500 ms (i.e., after

the first phase completed), the 18 objects located on the rings were presented. This was the start of the search phase and the phase concluded when the participant pressed either “j” or “f” on the keyboard. The target to find placed in the center of the screen was not redrawn nor removed at the start of the search phase. An example of this presentation is included in Figure 2.

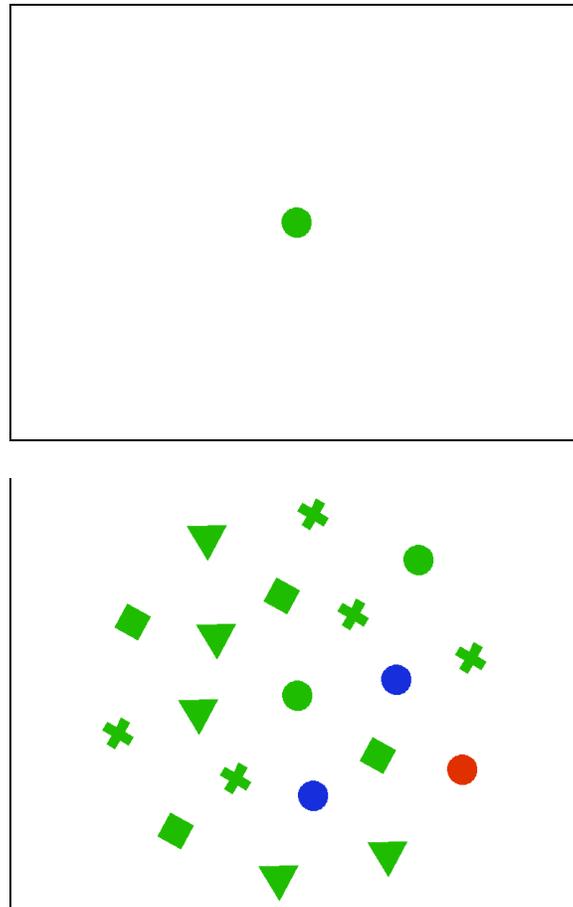


Figure 2. Example of the display phases for a target to find and then concentric ring “presentation”.

If the concentric rings were presented before the target to find, the 18 objects were displayed on the screen, and after the first phase completed, the target to find appeared in the center of the screen. This was the start of the search phase

and the phase again concluded when the participant pressed either “j” or “f” on the keyboard. The 18 objects on the rings were not redrawn nor removed at the start of the search phase. An example of this presentation is included in Figure 3.

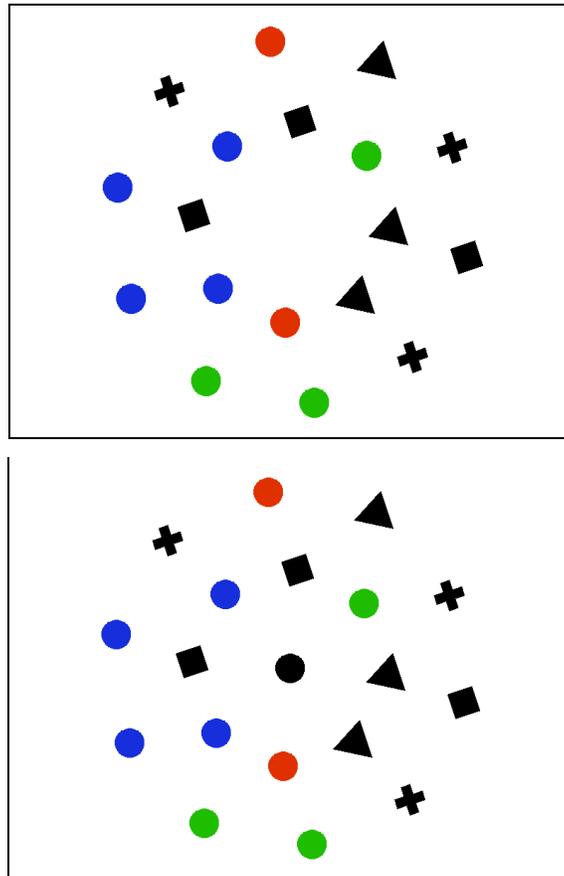


Figure 3. Example of the display phases for a concentric ring and then target to find “presentation”.

If the 18 objects on the concentric rings were shuffled whilst presenting the target to find, this occurred after the first phase completed. The task was therefore identical to the condition previously discussed except that when the target to find was presented, the 18 objects on the ring were removed, shuffled, and redrawn. An example of this presentation is included in Figure 4.

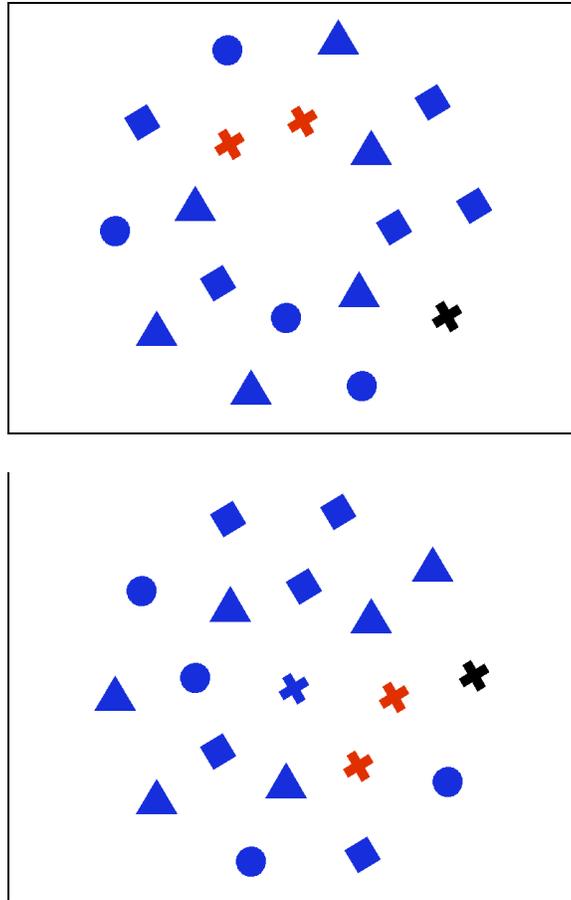


Figure 4. Example of the display phases for a concentric ring and then target to find whilst swapping “presentation”.

If a mistake was made (i.e., a false alarm or miss), the screen flashed for 3.2 seconds while disabling user input. This method aimed to keep the threshold for each subject performing the signal detection task relatively consistent while minimizing error rates and maximizing throughput.

Both eye-tracking and timing data were recorded. An experimenter monitored each experiment trial and recalibrated the eye tracker as needed. All trials were presented in one session. Subjects were asked to fill out a short survey

after completing the session. Subjects were also urged to use both hands while performing the experiment to minimize dispersion in motor response times.

Dependent Variables

Saccade landing point. Fixations were detected using the velocity-sensitive algorithm formalized by Salvucci and Goldberg (2000). Once the search task began (i.e., once all objects were drawn on the display), the next saccade's horizontal and vertical fixation points were recorded. The saccade landing point is defined as the ending location of this saccade; in other words, the ending location of the first saccade recorded after the search phase begins.

Calculating the projected landing point. Often, the landing point of the first saccade fell short of the inner ring or in between the inner and outer rings. If the target was present, and one assumes that the participant was initiating a saccade to land on the ring with the target, then the vector formed from the saccade's starting and ending point can be extended until it intercepts with this ring. This is analogous to drawing a line on the display that runs through this vector, and then recording the location where this line intercepts with the ring that the target was on.

Changing to polar coordinates. Because the task was designed with the objects arranged on concentric rings, saccade landing points were converted from Cartesian to polar coordinates. The angle of the saccade landing point, or "saccade direction", is the angle of the *projected* saccade landing point, while the radius of the saccade landing point, or "saccade distance", is simply the absolute distance between the *actual* saccade's starting and ending point.

For instance, a saccade initiated from the center of the display landing directly on the far-left object on the inner ring would have a saccade direction of 270 degrees and a saccade distance of 5.7 degrees (i.e., the eccentricity of the objects on the inner ring). On the other hand, a saccade initiated from the center of the display landing 1 degree short of the far-left object on the inner ring would still have a saccade direction of 270 degrees but a saccade distance of only 4.7 degrees.

Remapping saccade direction. All saccade directions were remapped so that 0 degrees represented a saccade directly on target while 180 degrees meant the saccade was initiated in the direction completely opposite the target (saccades landing to the right and left of the target were given positive and negative signs respectively). However, if 2 targets were displayed, the on-target location was the middle point between the targets. Therefore, with 2 spaced targets on the inner ring (i.e., 1 distractor in between), if a saccade lands on the left target, the direction would be $-1/8$ of 360, or -45 degrees.

Saccade directional error. Another useful variation to the saccade direction is the absolute directional error of the saccade to the closest object. For 1 target, this “directional error” is simply the absolute value of the saccade direction. However, when 2 targets are displayed, the directional error is calculated by first determining which target is closest to the saccade landing point and then calculating the absolute angular deviation between this target and that landing point. For instance, if 2 spaced targets are displayed on the inner ring at 0 and 90 degrees, and the projected saccade landing point is at 82 degrees, then the saccade directional error is 8 degrees.

Saccade latency. The time between the onset of the search task and the initiation of the first saccade is the “saccade latency.”

Results

Trimming

Practice trials. A short amount of time (usually 1-3 trials) was required for the participants to get completely comfortable with the task. To help minimize these practice effects, the first 5 trials for each participant were excluded.

Target absent trials. We wanted to test interactions between the salience of the target and the rest of the experimental manipulations (e.g., relative spacing, presentation). Therefore, all target absent trials were excluded.

Sporadic saccade. In certain cases, a saccade could be initiated that has no projected landing point (i.e., it does not intersect the ring with the target). For example, say that the saccade is initiated from the far left corner of the screen and lands on the far right corner. In this case, if the target is located on the inner ring, there is not enough information to calculate a projected landing point since the horizontal line formed at the top of the screen doesn't intersect the ring drawn in the middle of the screen. For these rare cases (~ .4 % of target-present trials), the saccade was labeled as “sporadic” and excluded.

Short saccade. Several cases arose where participants failed to initiate a saccade while performing the search task. First, a null zone was defined spanning 1.2 degrees radially outward from the center of the screen. If the landing point of the first detected saccade happened to fall anywhere inside this circular region, the saccade was excluded. This occurred for ~ 13% of target-present trials.

No saccade. Next, if the velocity-sensitive algorithm used to detect a saccade failed to register a single saccade during the trial (i.e., the eyes did not move during the trial), the saccade was labeled as “saccadeless.” How often this occurred was highly dependent on the particular subject. A few subjects were even able to keep their eyes focused on the center of the screen throughout the entire experiment, thereby labeling up to 95% of trials as saccadeless. When interviewed afterwards, these participants said they made an active choice not to move their eyes because they felt they could complete the task faster with this strategy (which they could, with ~ 14% decrease in target-present RTs compared to eye movers). For the purposes of this study, we were not interested in modeling these participants choosing to control their eye movements using high-level cognitive strategies. Therefore, we needed to pick only those participants that were allowing their eyes to move naturally around the display, which presumably were the participants with a small percentage of saccadeless trials.

Looking at Figure 5 (i.e., the distribution of percent saccadeless trials by participant), there is a clear grouping of subjects having less than or greater than 50% saccadeless trials.

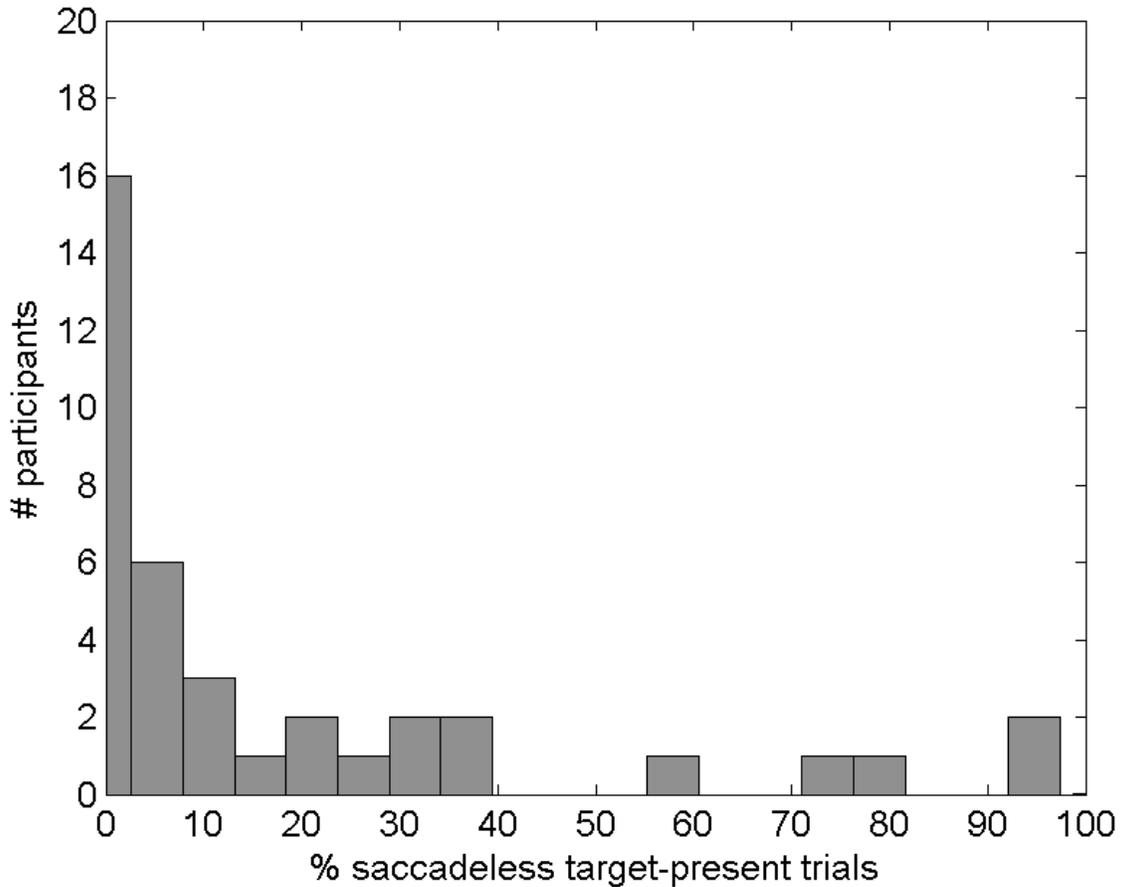


Figure 5. Distribution of percent saccadeless target-present trials by participant.

Further, since so many participants (i.e., 2/3) of the group having less than 50% had less than 10% saccadeless trials, this group was further subdivided at 10%. Subjects were therefore separated into 3 groups depending on the percent of saccadeless target-present trials ($\leq 10\%$, between 10% and 50% $\geq 50\%$). To be conservative, only subjects in the lower group (i.e., $\leq 10\%$ saccadeless trials) were kept, while the rest were removed from analysis (i.e., 13/38 subjects were removed). Of those subjects kept, $\sim 4\%$ of target-present trials were labeled as saccadeless and excluded.

Calibration. Several subjects' eyes could not be reliably tracked at all, so they were completely removed from analysis (i.e., 12 of 50 subjects were removed). Additionally, of the 38 subjects analyzed, ~ 2% of target-present trials had questionable calibration settings and were therefore excluded.

Outliers. Data greater than 3.5 standard deviations from the mean in either direction were excluded. This was done at the participant level and for 3 observed variables: response time, saccade latency, and saccade distance. In other words, if a trial included observed data more than 3.5 standard deviations from the participant's mean for *any* of these variables, the trial was excluded. ~ 2% of target-present trials were removed due to this procedure.

Remaining data. After all trimming, a total of 3540 of 6840 possible target-present trials, spanning across 25 subjects, were kept for further analysis. The average number of target-present trials used for each subject was therefore ~ 141 of the 180 available (i.e., ~ 22% of trials were excluded for each subject).

Although this is a fairly large amount of excluded trials per subject, the majority of these trials were cases where the first saccade landing point failed to span far enough away from the center of the display (i.e., the short saccades that occurred for ~ 13% of target-present trials). Presumably, these first saccades were made to encode the target to find, which means the next recorded saccades were made towards the first targeted object. However, we were testing how the age of the objects displayed influenced saccadic accuracy (i.e., how "presentation" influenced accuracy), and the age of the objects increases with each successive saccade (i.e., as the eyes continue to move around, time progresses, and the objects are displayed

longer). Therefore, we excluded these trials to ensure that the age of the objects displayed was *only* a function of the “presentation” manipulation, so that the effect of age on accuracy could be directly assessed by testing the effect of presentation on accuracy.

Distributions

Saccade distance. Included in Figure 6 is the distribution of observations for saccade distance. The mean and standard deviation were 3.18 and 1.70 degrees respectively.

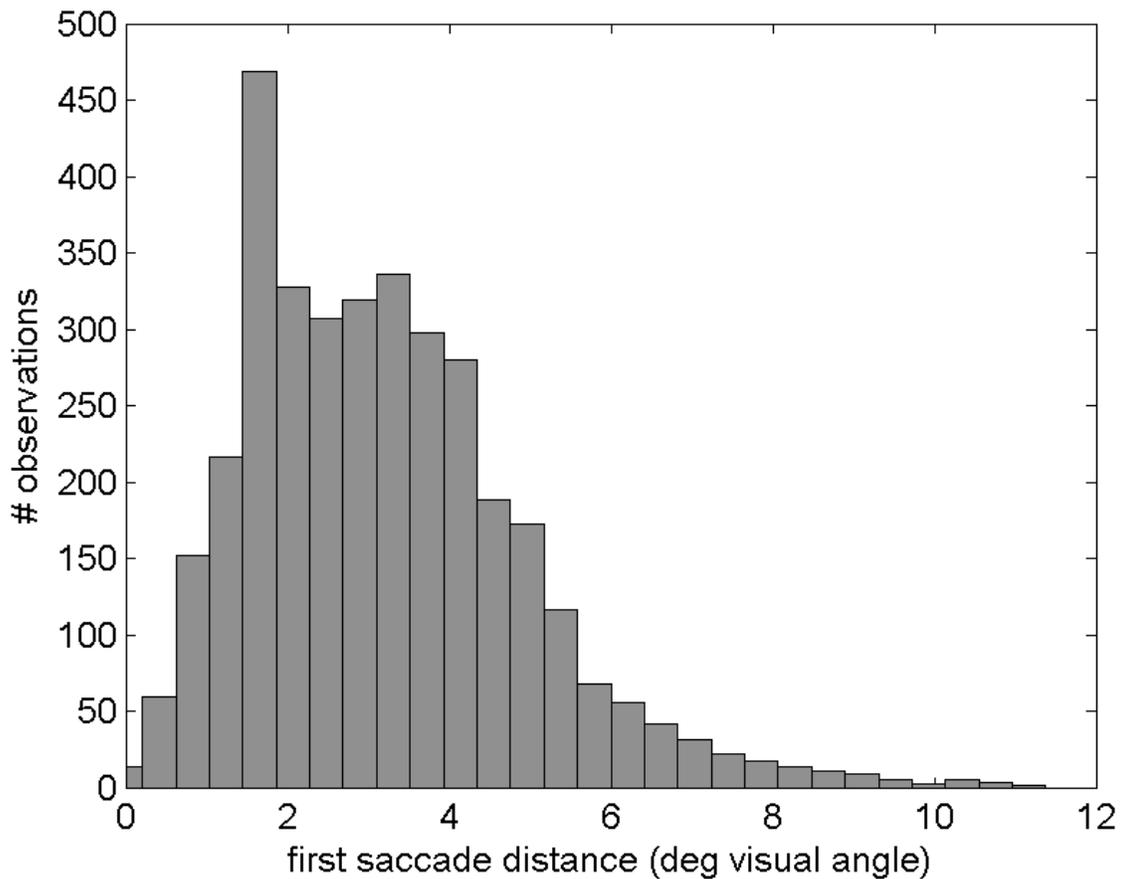


Figure 6. Distribution of observations for saccade distance.

Saccade direction. Included in Figure 7 is the distribution of observations for saccade direction. Note how the distribution is centered and peaks around 0 degrees. This makes sense qualitatively, for the participants are looking for the target, and the target is located at 0 degrees for the singleton condition, -45,45 degrees for 2 spaced targets on the inner ring, and -22.5,22.5 degrees for 2 grouped targets on the inner ring (symmetrical across 0 degrees for the outer ring condition as well).

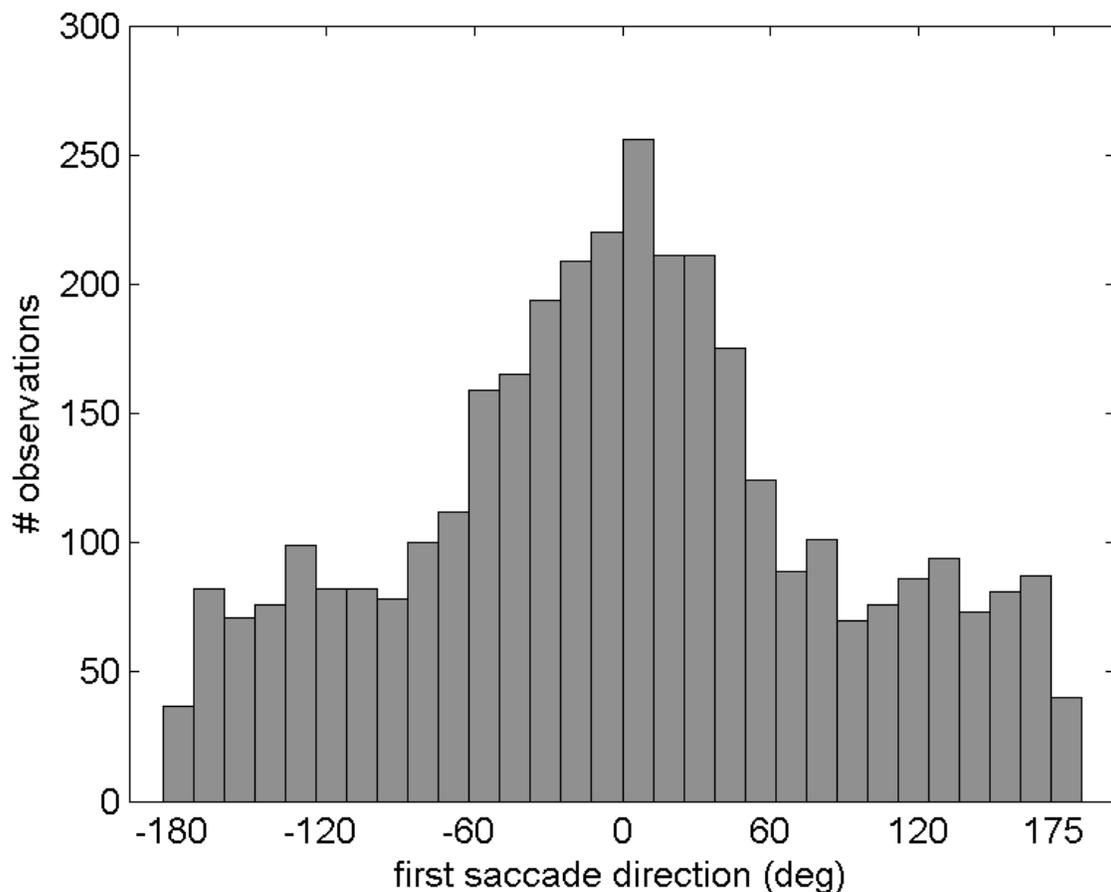


Figure 7. Distribution of observations for saccade direction.

Saccade directional error. Included in Figure 8 is the distribution of observations for saccade directional error. This is more or less the saccade direction distribution folded onto itself by making a crease at 0 degrees.

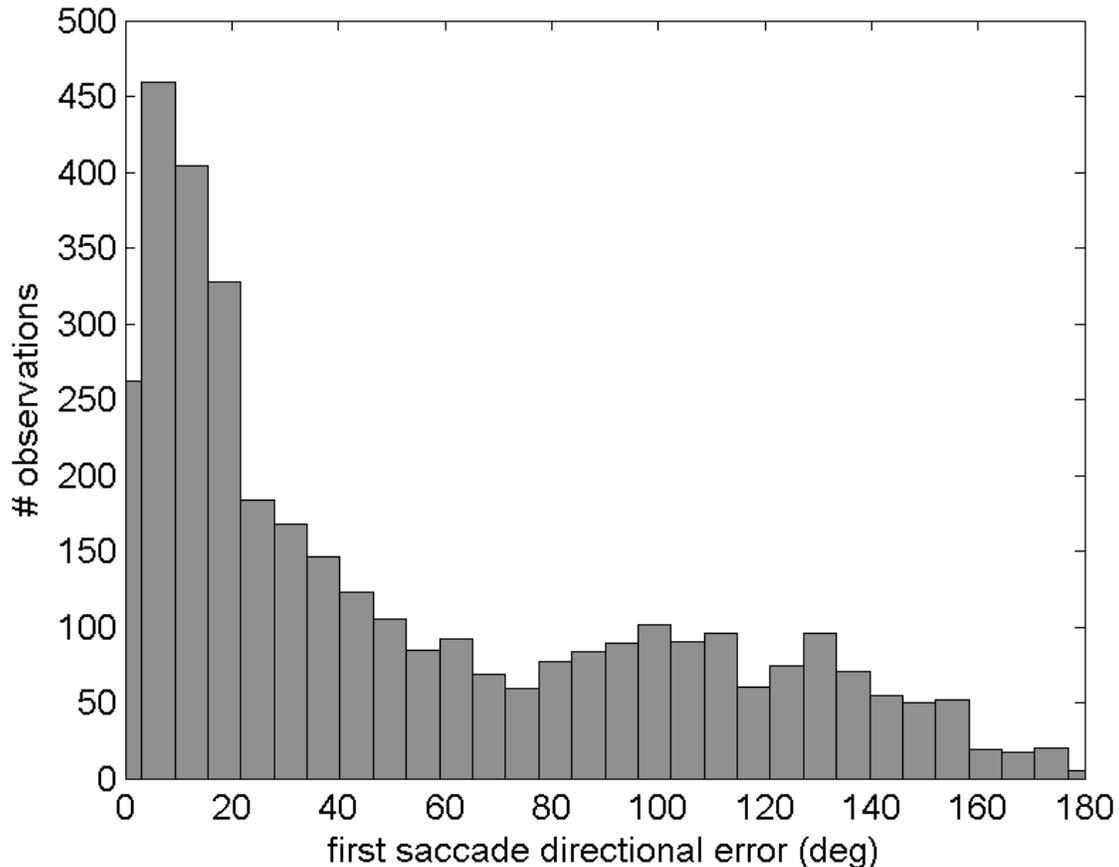


Figure 8. Distribution of observations for saccade directional error.

Saccade latency. Included in Figure 9 is the distribution of observations for saccade latency. These observations have been partitioned into 2 groups. The left group consists of observations where the target is presumably encoded before the search phase begins (i.e., the target to find and then concentric ring presentation; see Figure 2). The right group consists of observations where the target must be encoded after the search phase begins (i.e., the other 2 “presentations”; see Figures 3 and 4). Note the qualitative similarities between the left and right groups: the

large amount of dispersion in the distributions as well as the clustering of saccades with little or no latency (i.e., the 2nd peak of the distributions close to the origin). For completeness of modeling (i.e., to model the effect of latency on accuracy between the entire observed range), these short-latency saccades were not excluded from analysis. The mean and standard deviation for the left distribution were 302 and 142 ms respectively while the descriptives for the right distribution were 321 and 159 ms.

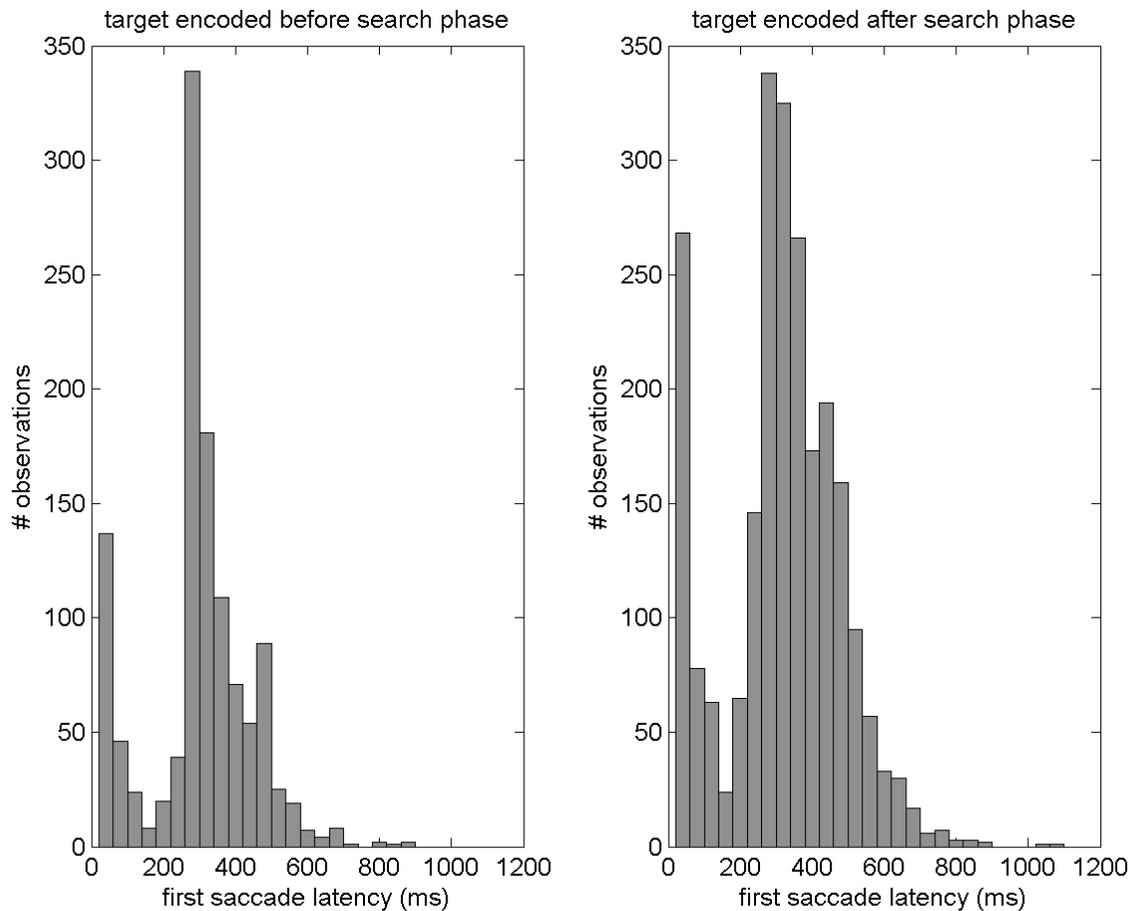


Figure 9. Distribution of observations for saccade latency.

Reorganizing Data

Task presentation. Due to the high similarity between the distributions of saccade latencies for the 3 task presentations ($r^2(1168) = .983$, $r^2(1182) = .989$, and $r^2(1168) = .989$ for the Q-Q plots crossed between the 3 presentations; see Figure 10), one might argue that task presentation is merely influencing the likelihood that the target to find will be encoded before the search phase begins.

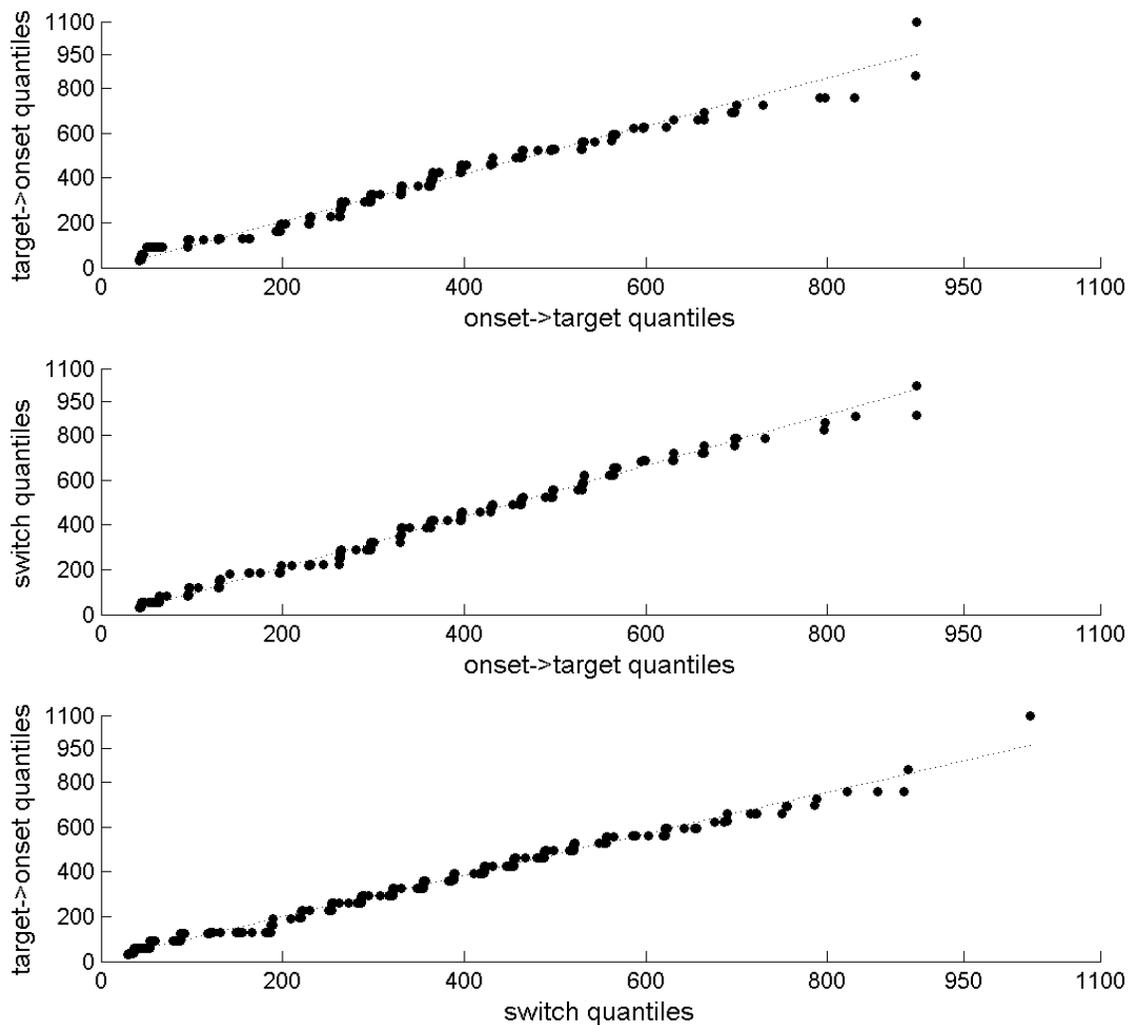


Figure 10. Q-Q plots crossed between the 3 presentations .

In this case, the mean difference between the left and right groups in Figure 9 represents roughly the amount of time required to encode the target to find. If that amount of time is subtracted off the saccade latencies for observations where the target must be encoded after the search phase begins (i.e., the right group in Figure 9), what remains is the distribution of saccade latencies *after* the target to find has been encoded. This distribution should line up with the left group in Figure 9, assuming that participants are encoding the target to find as soon as it is presented on the display.

Using this method, an additional variable was derived for analysis: the *adjusted* saccade latency. For trials where the target must be encoded after the search phase begins (i.e., the right group in Figure 9), the adjusted saccade latency is the observed saccade latency minus the average amount of time required to encode the target to find (25 ms was used for every subject). For trials where the target to find is encoded before the search phase begins (i.e., the left group in Figure 9), the adjusted saccade latency is simply the observed saccade latency.

Binning the salience of the target. The ratio of color/shape distractors affects the signal to noise ratio of the target relative to the distractors (i.e., the salience of the target). The mean observed response times for cases with low, medium, and high numbers of same-color distractors were 982 ($CI_{.95} = 951,1013$), 1021 ($CI_{.95} = 995,1046$), and 1029 ($CI_{.95} = 1001,1056$) ms respectively. This suggests that participants are primarily guided by color when finding the target (means monotonically increasing from left to right), but partially guided by shape as well (rate of increase in means decreases from left to right). Byrne's theory of visual

saliency (2006) argues that search time is inversely related to the relative saliency between the target and distractor (i.e., high signal to noise ratio leads to efficient searches while low signal to noise ratio leads to inefficient searches). Because response times are roughly the same for conditions with medium and high numbers of same-color distractors, and substantially shorter with low numbers of same-color distractors, one might argue that target saliency for the first two conditions is roughly equal and relatively low, while target saliency for the latter condition is substantially larger. Thus, trials with low numbers of same-color distractors were labeled as having a “high” target saliency, while trials with medium and high numbers of same-color distractors were labeled as having a “low” target saliency.

Binning the ring. Mean observed response times were largely influenced by ring (906 ($CI_{.95} = 885,927$) and 1109 ($CI_{.95} = 1085,1133$) ms for the inner and outer ring respectively), but only marginally influenced by the number and relative spacing of targets within a ring (952 ($CI_{.95} = 890,1005$), 908 ($CI_{.95} = 874,943$), and 886 ($CI_{.95} = 858,916$) ms for inner singleton, inner spaced, and inner grouped respectively; 1175 ($CI_{.95} = 1124,1226$), 1103 ($CI_{.95} = 1062,1143$), and 1079 ($CI_{.95} = 1043,1115$) ms for outer singleton, outer spaced, and outer grouped respectively). Therefore, to increase power when running the inferential statistics, the number and relative spacing of targets were stripped and only the ring of the target was coded.

Binning the adjusted saccade latency. To help qualitatively interpret the large effect of saccade latency on the observed variables, an additional variable was created, which partitions observed saccade latencies into 3 groups: short, medium,

and long. This was done at the participant level separating the saccade latencies into 3 regions with equal numbers of observations in each region (i.e., a 2-median split). However, the inferential statistics were *not* tested using these bins, for doing so would cause information to be lost across the saccade latency dimension. This technique was therefore used for qualitative purposes only (e.g., cleaner visualizations).

Effects of Experimental Manipulations

Inferential technique. Because of the substantial amount of trimming performed, as well as the unequal binning of trials across the distractor ratio dimension, the number of observations in each cell for this design was highly unbalanced (these values are included in Table 2 for reference).

Table 2. Number of observations in each cell of the experimental design.

Salience	Ring	Presentation	n
High	Inner	Onset->target	190
High	Inner	Switch	188
High	Inner	Target->onset	181
High	Outer	Onset->target	200
High	Outer	Switch	202
High	Outer	Target->onset	202
Low	Inner	Onset->target	379
Low	Inner	Switch	385
Low	Inner	Target->onset	390
Low	Outer	Onset->target	417
Low	Outer	Switch	409
Low	Outer	Target->onset	397

Therefore, in order to correctly analyze the results in this dataset, the Restricted Maximum Likelihood (REML) method was used to estimate the variance components. This method has the appealing property in that the results obtained in a balanced design are mathematically equivalent to the results found with the more

traditional ANOVA. However, the REML method is useful for much broader applications; specifically, handling truncated and missing data in mixed designs (Dempster, Laird, & Rubin, 1977).

Dependent variables tested. Saccade direction is not the appropriate dependent variable to use in this situation, for the expected mean of the saccade direction will always be zero regardless of the targeting accuracy of the first saccade. For example, if the target is no more salient than the rest of the distractors, the distribution of saccade direction should be a continuous uniform with a mean of zero (saccade accuracy is at chance levels). On the other hand, if the target is much more salient than the rest of the distractors, the distribution of saccade direction should show a sharp peak at zero but still have a mean of zero. What has changed between these situations is not the mean but the dispersion of saccades around the target location (smaller dispersion means more accurate targeting).

The saccade directional error has the useful property that the mean directional error varies inversely with targeting accuracy (less accuracy leads to more dispersion which leads to a higher mean absolute deviation from the target). However, because the *absolute* deviation from the target is used to calculate directional error, the statistic is no longer normally distributed. Nevertheless, the REML statistical technique has been shown to be fairly robust to violations of normality assumptions (e.g., Banks, Mao, & Walter, 1985). Therefore, saccade directional error, along with saccade distance, were used as the two tested dependent variables. The effects of the experimental manipulations on these dependent variables using REML are summarized in Tables 3 and 4.

Table 3. Effect of experimental manipulations on saccade distance.

Fixed Effect	DF Num	DF Den	F	p
Ring	1	28.66	54.26	<.001
Saliency	1	15.06	.897	.359
Ring * Saliency	1	29.57	5.00	.033
Presentation	2	49.31	3.67	.033
Ring * Presentation	2	54.53	.926	.402
Saliency * Presentation	2	50.31	.901	.413
Ring * Saliency * Presentation	2	29.68	.533	.592
Latency	1	25.92	142.14	<.001
Ring * Latency	1	19.14	26.30	<.001
Saliency * Latency	1	5.45	.398	.554
Ring * Saliency * Latency	1	1461	2.86	.091
Presentation * Latency	2	48.24	2.09	.135
Ring * Presentation * Latency	2	57.55	1.81	.173
Saliency * Presentation * Latency	2	650.4	0.00	1.000
Ring * Saliency * Presentation * Latency	2	5.39	.720	.528

Note: Bold denotes $p < .05$.

Table 4. Effect of experimental manipulations on saccade directional error.

Fixed Effect	DF Num	DF Den	F	p
Ring	1	36.68	54.79	<.001
Saliency	1	28.53	14.90	<.001
Ring * Saliency	1	6.64	10.09	.017
Presentation	2	52.06	4.18	.021
Ring * Presentation	2	52.42	.038	.963
Saliency * Presentation	2	58.59	1.12	.334
Ring * Saliency * Presentation	2	3379	2.23	.107
Latency	1	18.65	120.53	<.001
Ring * Latency	1	23.75	2.33	.141
Saliency * Latency	1	24.97	1.18	.287
Ring * Saliency * Latency	1	2.69	13.23	.043
Presentation * Latency	2	21.66	.471	.631
Ring * Presentation * Latency	2	50.3	1.48	.238
Saliency * Presentation * Latency	2	52.29	.523	.596
Ring * Saliency * Presentation * Latency	2	36.29	1.26	.296

Note: Bold denotes $p < .05$.

Effect of adjusted latency. The adjusted saccade latency accounts for the largest amount of variance in the model regressed on both dependent variables ($F(1,25.9) = 142.14, p < .001$ and $F(1,18.7) = 120.53, p < .001$ for saccade distance

and saccade directional error respectively), and a substantial interaction between adjusted saccade latency and ring is driving the main effect of latency on saccade distance ($F(1,19.14) = 26.30, p < .001$). This interaction is plotted in Figure 11. Note particularly how the slope for the effect of saccade latency on distance varies with ring (outer ring slope is roughly twice that of the inner ring). This makes sense intuitively, for objects on the inner ring are located at roughly half the eccentricity of objects on the outer ring (5.7 and 10.2 degrees for the inner and outer ring respectively). However, note how the mean saccade distance falls well short of the target object eccentricity, even when the target object is placed on the inner ring.

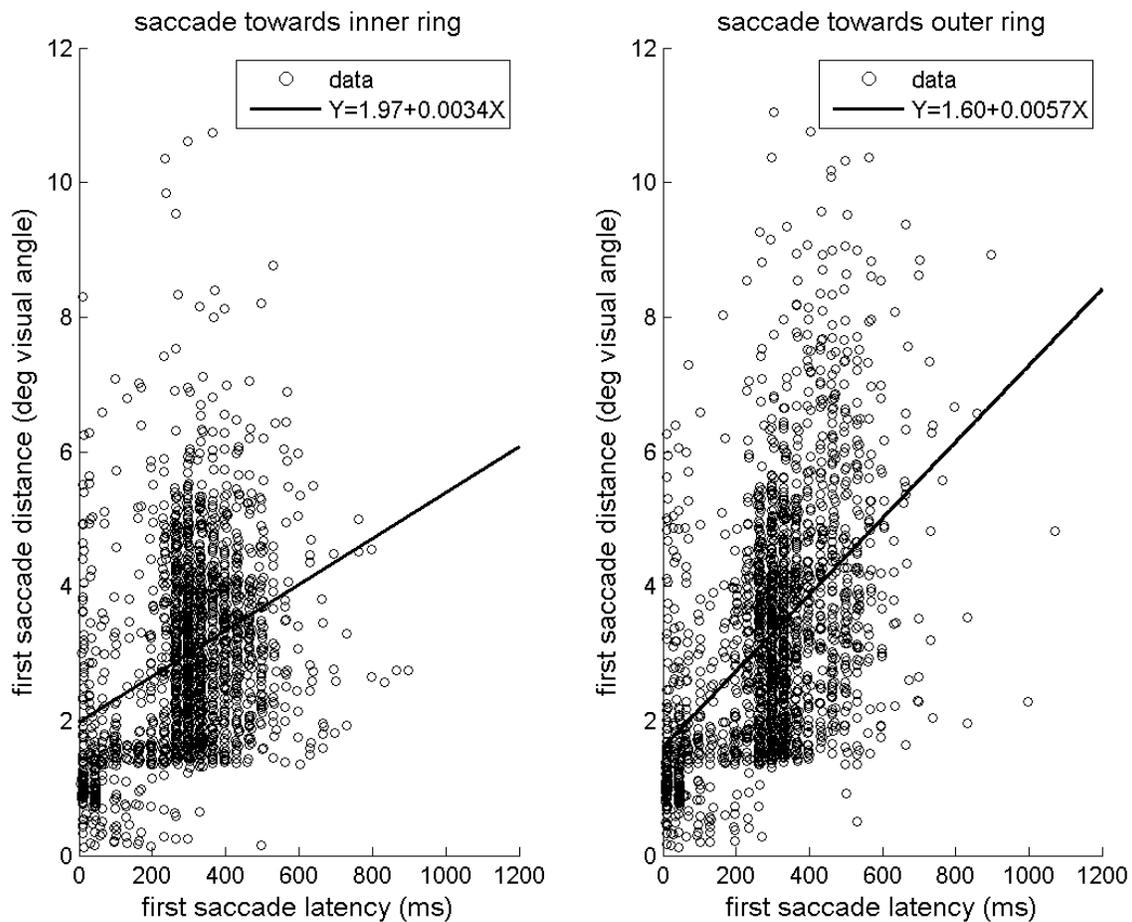


Figure 11. Effect of ring and adjusted saccade latency on saccade distance.

To better visualize how saccade latency influences saccade directional error, binned distributions of saccade direction for short, medium, and long saccade latencies are included in Figures 12 and 13. First note the speed/accuracy tradeoff evident in these plots (as saccade latency increases, the saccade is more likely to be towards a target object). Next, note that when 2 targets are grouped together (i.e., no distractor in between), the longest saccade latencies still show overlapping distributions between the 2 targets. However, when 2 targets are spaced apart (i.e., one distractor in between), the longest saccade latencies clearly show separate distributions for each target. Finally, note how targeting accuracy is just a bit sharper when the targets are placed on the inner ring compared to when they're on the outer ring.

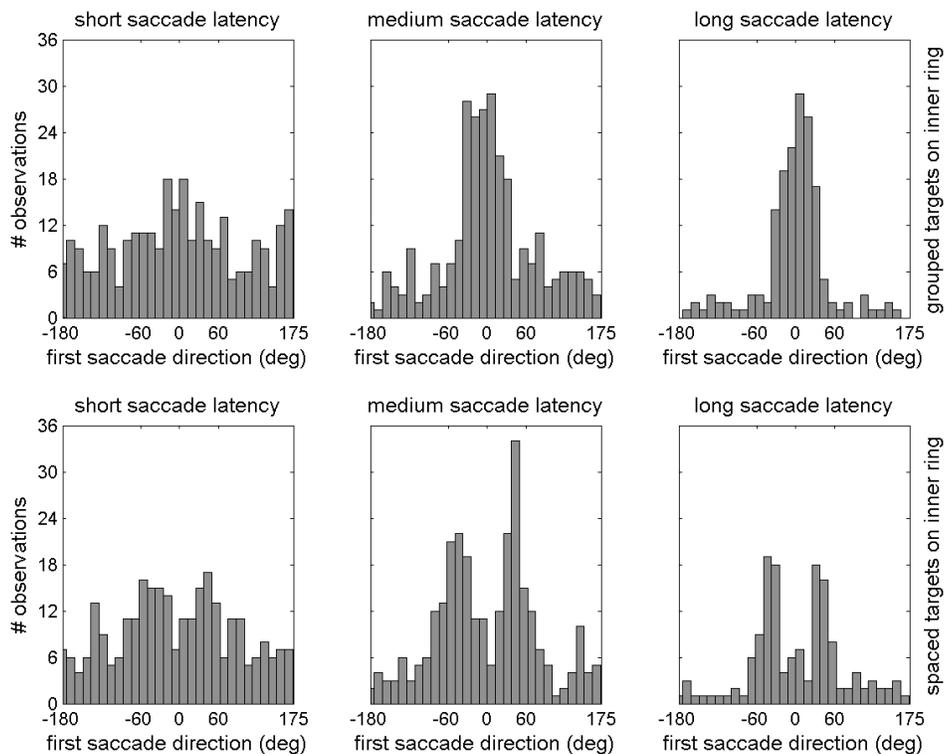


Figure 12. Histogram of saccade direction vs. latency bin for targets on inner ring.

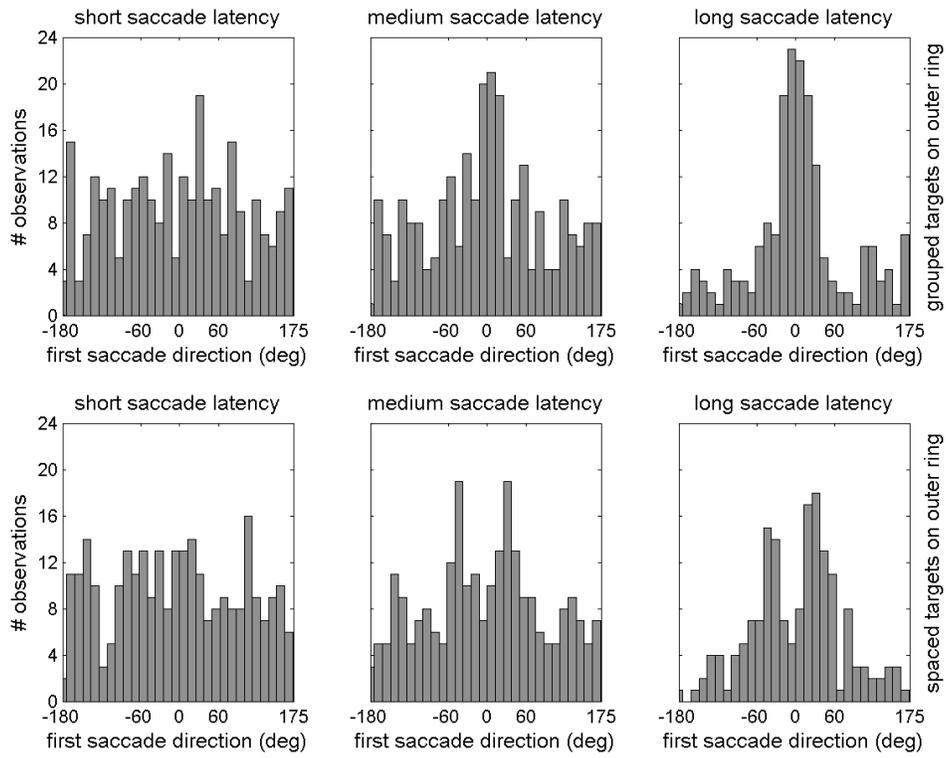


Figure 13. Histogram of saccade direction vs. latency bin for targets on outer ring.

Effect of salience and ring on saccade distance. The salience of the target did not substantially influence the saccade distance ($F(1,15.06) = .90, p = .36$; see Figure 14). However, the interaction between salience and ring did have a strong influence on saccade distance ($F(1,29.57) = 5.00, p = .033$; see Figure 15), and the main effect of ring on saccade distance was present as well ($F(1,28.66) = 54.26, p < .001$; see Figure 16).

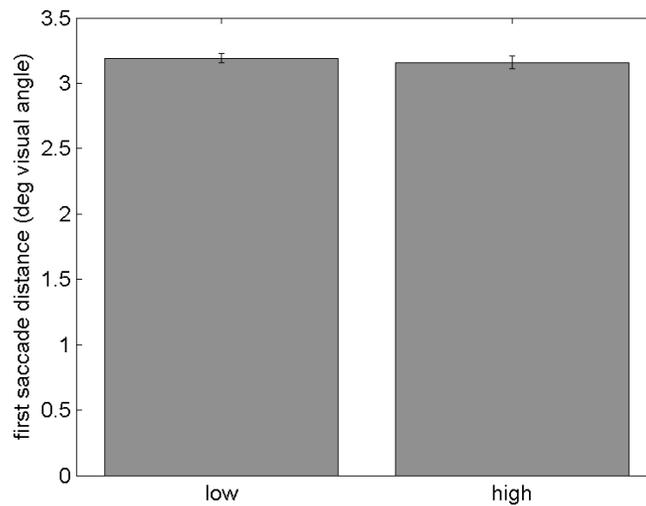


Figure 14. Observed means for the null effect of salience on saccade distance.

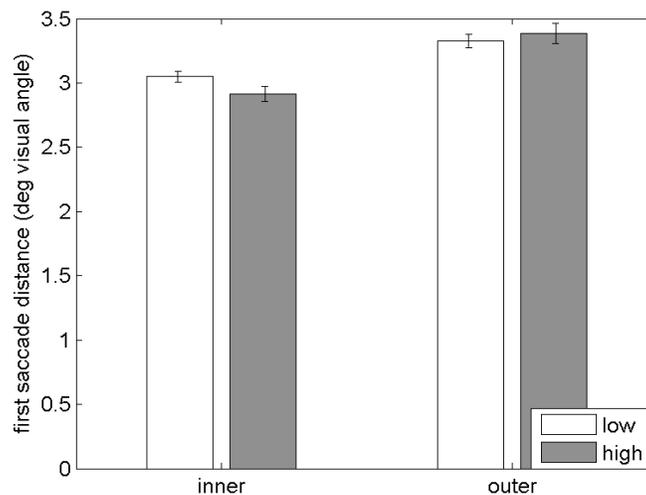


Figure 15. Observed means for the interaction between salience and ring on saccade distance.

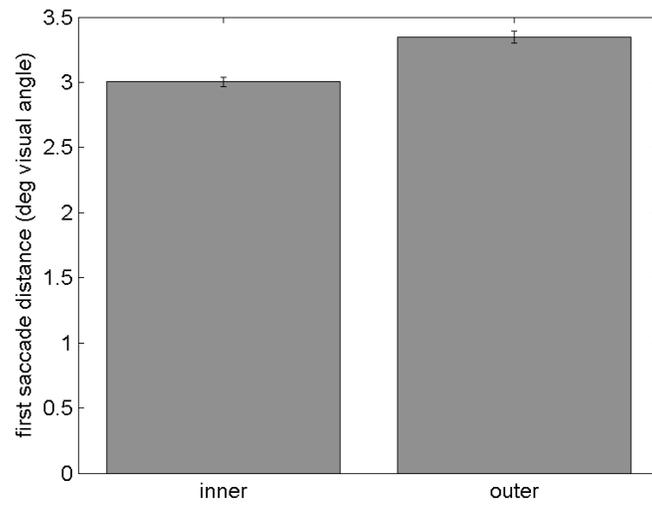


Figure 16. Observed means for the effect of ring on saccade distance.

Effect of salience and ring on saccade directional error. The salience of the target affected the targeting accuracy of the saccade ($F(1,28.53) = 14.90, p < .001$; see Figure 17), and an interaction between ring and salience on directional error was present as well ($F(1,6.64) = 10.09, p = .017$; see Figure 18). Finally, saccade targeting was more accurate overall when the target object was located on the inner ring compared to the outer ring ($F(1,36.68) = 54.79, p < .001$; see Figure 19).

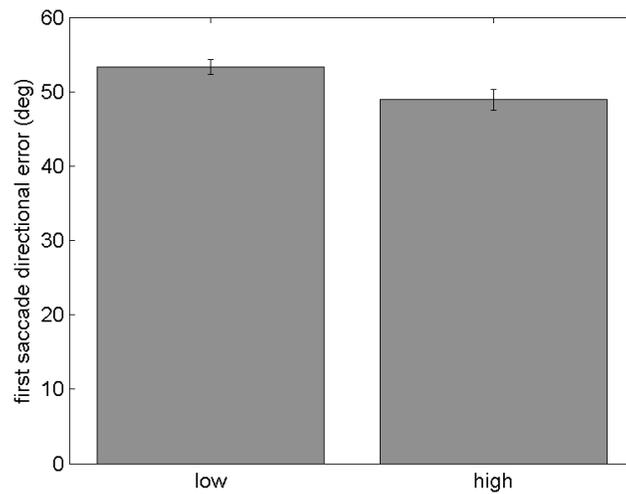


Figure 17. Observed means for the effect of salience on saccade directional error.

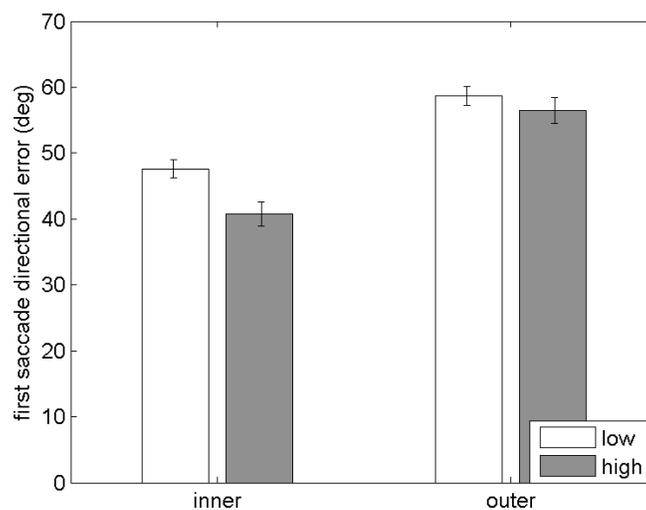


Figure 18. Observed means for the interaction between salience and ring on saccade directional error.

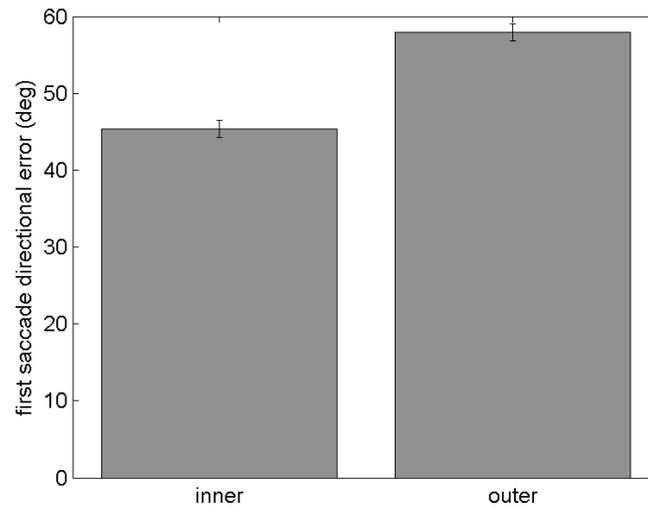


Figure 19. Observed means for the effect of ring on saccade directional error.

Effect of presentation. After adjusting the saccade latency, task presentation (see Figures 2, 3, and 4) still influenced saccade direction and directional error ($F(2,49.31) = 3.67, p = .033$ and $F(2,52.06) = 4.18, p = .021$ respectively; see Figures 20 and 21 respectively).

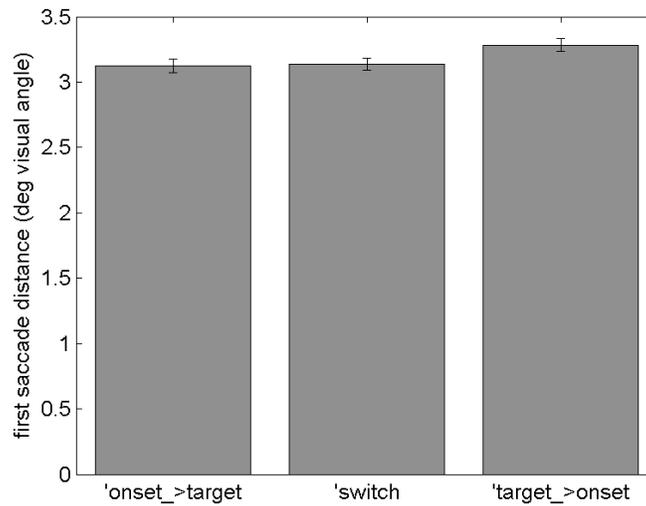


Figure 20. Observed means for the effect of presentation on saccade distance.

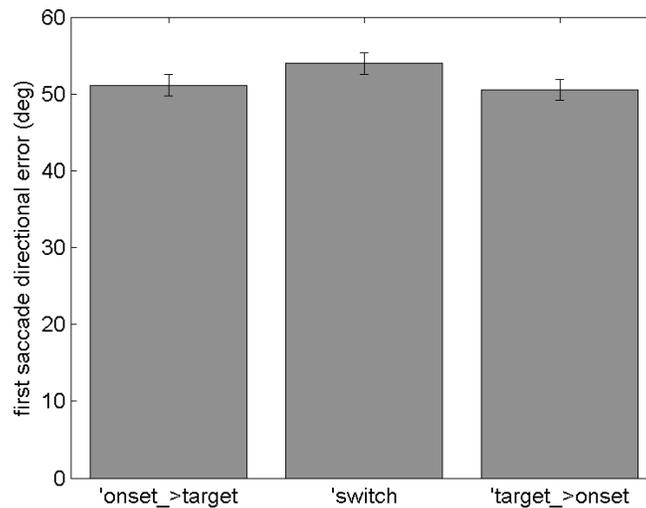


Figure 21. Observed means for the effect of presentation on saccade directional error.

MODEL

Towards Modeling Effects

The two hypotheses discussed (i.e., the speed/accuracy tradeoff residing in the salience system or targeting system) were considered to model the experimental results. However, the speed/accuracy tradeoff residing in the salience system was ruled out due to inconsistent predictions between the model and observed data, particularly across strong effects.

H[1]: The Speed/Accuracy Tradeoff Resides in the Salience System

Assume that the majority of the temporal dynamics reside in the system that determines the salience of each object. This salience map takes about 400 ms to evolve (Rao et al., 2002) and the landing point of the saccade is determined by taking the vector average of the salience map at the time the saccade is initiated. Quite an interesting finding for this study was the minimal effect task presentation had on saccade distance and targeting accuracy. If the salience map was evolving during the first 400 ms, a case where the objects on the concentric rings are displayed before the search task begins (and not swapped when the search task begins), should show high targeting accuracy since the salience map has already converged to its resting value. Conversely, targeting accuracy should decrease dramatically in cases where the eye is moved less than 400 ms after the concentric rings are presented (basically all data collected for the presentation case where the target to find is encoded before the search task begins).

Although the effect of presentation on directional error was statistically reliable ($F(2,52.06) = 4.18, p = .021$), the mean values for the cases where the

concentric rings are displayed ~ 1000 ms before and immediately when the search phase begins were almost identical (50.5 ($CI_{.95} = 47.8, 53.2$) and 51.1 ($CI_{.95} = 48.4, 53.9$) degrees respectively), and the case where objects on the concentric rings were switched turns out to drive the effect (53.9 ($CI_{.95} = 51.2, 56.7$) degrees). Since having the search objects displayed for ~ 1000 ms before starting the search task fails to increase targeting accuracy dramatically, the salience system must be evolving extremely quickly. Further, since the targeting system has a substantial temporal component (latency strongly influences saccade distance and accuracy), and the salience system is evolving at a much faster rate, the targeting system cannot be using the salience system to drive this speed/accuracy tradeoff.

One might argue that the reason for this null difference is because the search objects being used are too simple. It would certainly be interesting to see if the null effect still occurs using more complicated search objects (similar to those in Rao et al.'s study), for if a large effect is still missing, it would make a much stronger case refuting the hypothesis that the salience system is performing the vector averaging.

H[2]: The Speed/Accuracy Tradeoff Resides in the Targeting System

An alternative hypothesis (the one used in this study) is that calculation of each object's salience occurs almost instantaneously, or at least that the time required is too small to account for the fact that latency was still influencing accuracy 400 ms after the search task began (see Figures 12 and 13). Instead, a time-dependent targeting network controls the landing location of the saccade. This location is calculated by taking the vector average of the activations of all locations on this targeting network at a given time (each location corresponds to an object on

the display). Before a target is chosen, the activation of all locations more or less consists of transient noise, meaning that if the saccade were initiated at this time, the average landing location would be the same as the starting location. After an object is chosen however, the region in the targeting network corresponding to the location of that object will increase in activation, while regions corresponding to locations of other objects on the display will decrease in activation. The longer the latency (i.e., the more time that elapses), the greater the signal to noise ratio of the targeting system, the more accurate the targeting system becomes, and the closer the system will program an eye movement to land on the selected object.

Finally, the system is inherently noisy, and therefore the time when a “snapshot” of the activation in this network is taken and the saccade trajectory is calculated has some variance. In Table 5, note how only a few experimental variables affected adjusted saccade latency (most notably salience).

Table 5. Effect of experimental manipulations on adjusted saccade latency.

Fixed Effect	DF Num	DF Den	F	p
Ring	1	23.76	1.54	.227
Salience	1	21.67	10.05	.005
Ring * Salience	1	25.82	.480	.495
Presentation	2	53	1.12	.335
Ring * Presentation	2	58.49	2.05	.138
Salience * Presentation	2	47.63	3.91	.027
Ring * Salience * Presentation	2	60.26	.645	.528

Note: Bold denotes $p < .05$.

Further, even the largest effect, the salience manipulation, only changes the average saccade latency by ~ 11 ms (~ 302 ($CI_{.95} = 296,309$) and 289 ($CI_{.95} = 281,297$) ms for low and high salience conditions respectively). Therefore, the majority of variance present in saccade latency is not being caused by external manipulations

and instead must be due to internal noise inherent in the system. This maps nicely to the EMMA model, where the variance of time required to prepare the saccade is independent of external factors (e.g., target salience) and sampled from a gamma distribution (Salvucci, 2001).

Predictions: Effect of salience on saccade distance. Take 2 scenarios: [1] the target is highly salient, and [2] the target's salience is approximately equal to the rest of the objects on the display.

For [1], targeting accuracy is near perfect; therefore, times when the target is on the inner ring the eye will move to the inner ring, and times when the target is on the outer ring, the eye will move to the outer ring. Since half the trials have the target on the inner ring, and half have it on the outer ring, half the trials the eye will move to the inner ring and the other half to the outer ring.

For [2], targeting accuracy is near chance. Therefore, for [2] as well, half of the trials the eye will move to the inner ring, and half the trials to the outer ring.

This means that the average saccade distance should be roughly independent of target salience, thus predicting the null effect observed ($F(1,15.06) = .90, p = .36$). However, when analyzed by ring, an interaction between ring and salience on saccade distance should be present. To see this, note that the difference between [1] and [2] is not the average saccade distance but the average distance when the target is on a particular ring. Take [1], for example. When the target is on the inner ring, the saccade will be initiated to the inner ring, and vice-versa for the outer ring. This means the average saccade distance for the inner ring should be less than that for the outer ring. However, for [2], since everything is at chance levels, the average

saccade distance when the target is on the inner ring will be the *same* as that of the outer ring. Higher salience therefore tends to separate the means of saccade distance (target on outer and inner ring) apart from each other. This predicts a directional change in the effect of salience on saccade distance between the inner and outer ring (an interaction), which is precisely what is observed ($F(1,29.57) = 5.00, p = .033$).

Predictions: Effect of latency on saccade distance. Having a positive slope for saccade distance vs. latency in both cases (i.e., when the object is on the inner and outer ring) is predicted by H[2] as well. To see this, assume that the visual system is targeting an object on the inner ring. If the saccade is initiated with a short latency, the majority of the other objects on the display will still have a bit of activation associated with them in the targeting system, and therefore those objects will pull the vector average closer to themselves and collectively towards the origin. However, if the saccade is initiated with a relatively long latency, the signal to noise ratio of the targeted object compared to the rest of the objects will be much higher, meaning that the rest of the objects get less of a pull on the vector average towards their location, causing the eye movement to move outwards from the origin and closer to the targeted object. Thus, there should be a strong effect of latency on saccade distance. Analogous reasoning can be used for the case where the targeted object is on the outer ring; the only difference being that the slope should be larger since the targeting system has farther to go.

Predictions: The rest. All other manipulations should fit with H[2], except for task presentation, since H[2] does not handle anything concerning how old the

objects on the display are. The effects of presentation on saccade distance and directional error are most likely due to some of the salience temporal dependencies described in H[1]. However, it is important to emphasize that the age of the objects (i.e., the manipulations in presentation) do not drastically influence the first saccade distance (mean saccade distances of 3.13 ($CI_{.95} = 3.03, 3.22$), 3.29 ($CI_{.95} = 3.19, 3.39$), and 3.14 ($CI_{.95} = 3.05, 3.23$) degrees for the cases in Figures 2, 3, and 4 respectively), nor do they drastically influence the first saccade directional error (mean directional errors of 51.1 ($CI_{.95} = 48.4, 53.9$), 50.5 ($CI_{.95} = 47.8, 53.2$), and 53.9 ($CI_{.95} = 51.2, 56.7$) degrees for the same cases respectively). Additionally, RTs were not largely influenced by presentation either (983 ($CI_{.95} = 958, 1009$), 986 ($CI_{.95} = 959, 1013$), and 1063 ($CI_{.95} = 1032, 1094$) ms for those cases respectively). Therefore, although H[1] may be incorporated alongside H[2] to additionally account for the effects of presentation, it was not pursued here, for it was felt that capturing the majority of the variance with a simpler model was more informative than explaining those last few pieces with relatively small effects by adding additional model components.

The modeling work thus aimed to capture the majority of the variance observed by accounting for the effects of ring, salience, and adjusted latency on the dependent variables. However, it did not attempt to account for the remaining effect of task presentation. Therefore, after adjusting the saccade latency, all observations were collapsed across this presentation dimension, and the model was fit to these means.

Formalizing the Model: The Saccade Targeting System

Assume we have N objects arranged at various locations on the display. Each object i has an associated (x_i, y_i) location. Because the targeting system is noisy in determining the precise location of each object, some uncertainty is added when calculating each location (using the EMMA model, this noise term is sampled from a normal distribution with a standard deviation equal to $1/3$ the object's polar distance from the fovea). After adding noise, each object's location is defined as

$$X_i(r) = x_i + \varepsilon(r) \quad Y_i(r) = y_i + \varepsilon(r) \quad (3)$$

where capital terms are values after noise is added, lower-case terms are true values, and functional parenthesis imply a dependency on other values (specifically polar distance of the object for the noise term). This polar distance is defined as:

$$r_i = \sqrt{x_i^2 + y_i^2} \quad (4)$$

The targeting system takes the vector average of the location of each object on the display to calculate the saccade landing point. This vector average is weighted by the normalized targeting activation of each object in the targeting system

$$X(r, t) = \sum_{i=1}^N X_i(r) P_i(t) \quad Y(r, t) = \sum_{i=1}^N Y_i(r) P_i(t) \quad (5)$$

where the normalized targeting activation of a given object on the display is defined as:

$$P_i(t) = \frac{A_i(t)}{\sum_{i=1}^N A_i(t)} \quad (6)$$

The targeting activation of all objects on the display start with some amount of transient noise. This initial activation is defined as:

$$(A_o)_i = (a_o)_i + \varepsilon[(a_o)_i] \quad (7)$$

where $(a_o)_i$ is the initial activation of object i before noise is added, and ε is sampled from a Gamma distribution with s equal to the initial activation (quite noisy).

After the targeted object is chosen, its activation starts to increase while the rest of the objects' activation decreases. As time moves forward (i.e., as saccade latency increases), the signal to noise ratio of the targeted object's activation compared to the rest of the objects increases, causing X and Y of the targeting system to be weighted more and more by the targeted object. Rate parameters, similar to those used by Logan (1996), were implemented to formalize this temporal dependency of activation (albeit in a slightly different manner than Logan originally used them). The activation of an object in the targeting system is therefore defined as:

$$A_i(t) = (A_f)_i + [(A_o)_i - (A_f)_i] \cdot e^{-kt} \quad (8)$$

where the f and o subscripts define the final and initial activation of the object respectively, and k represents the rate parameter. The rate parameter for both targeted and non-targeted objects on the display was set to $.01 \text{ ms}^{-1}$ (i.e., having a time constant of 100 ms).

The final activation of an object depends on if that object is being targeted or not, and therefore $(A_f)_i$ will have 2 different values:

$$(A_f)_T = (A_o)_T \cdot g_T \quad (A_f)_{NT} = (A_o)_{NT} \cdot g_{NT} \quad (9)$$

Here, T and NT represent the targeted and non-targeted objects respectively, while g corresponds to the gain assigned to these types of objects. The gain terms used to

fit these data were 2.5 and .125 for the targeted and non-targeted objects respectively (a zero gain for the non-targeted objects was considered but did not quite fit the data as accurately as a non-zero gain).

Finally, the initial activation of an object was set to 1 for all objects on the display:

$$(a_o)_i = (a_o)_T = (a_o)_{NT} = 1 \quad (10)$$

This is actually not a free parameter, for if this value is multiplied by some factor, the gain coefficients can just be multiplied by that same factor to produce identical results (a summary of the free parameters used for the targeting system is included in Table 6).

Table 6. Free parameters used in the targeting system to model data.

Parameter Name	Parameter Symbol	Parameter Value
Location sigma	s_l	1/3 the mean
Activation sigma	s_a	1 x the mean
Activation Rate	k	.01 s ⁻¹
Target Gain	g_T	2.5
Non-target Gain	g_{NT}	.125

Visualizing the model's performance. A Monte Carlo simulation ($N = 50,000$) using this targeting system was run for the two ring cases (i.e., where the target object is located on the inner and outer ring). Figures 22 and 23 show the results of this simulation using contour plots of the probability density functions (across a range of saccade latencies) for saccade distance and saccade direction respectively. Each plot breaks out the two ring cases, and shows how latency strongly influences both the accuracy and distance of the saccade.

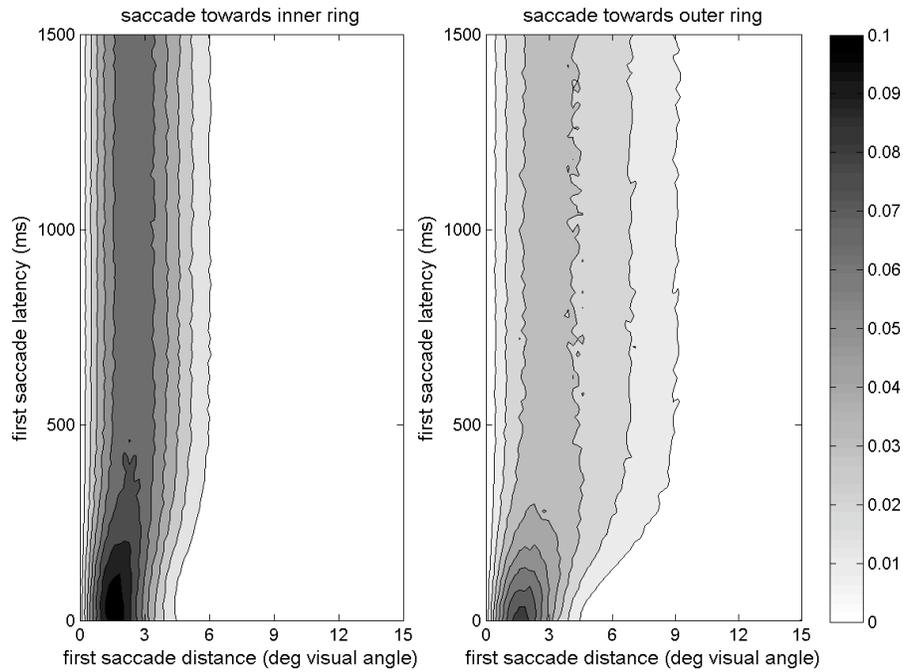


Figure 22. Contour plot of the targeting system's predicted saccade distance as a function of saccade latency and targeted ring.

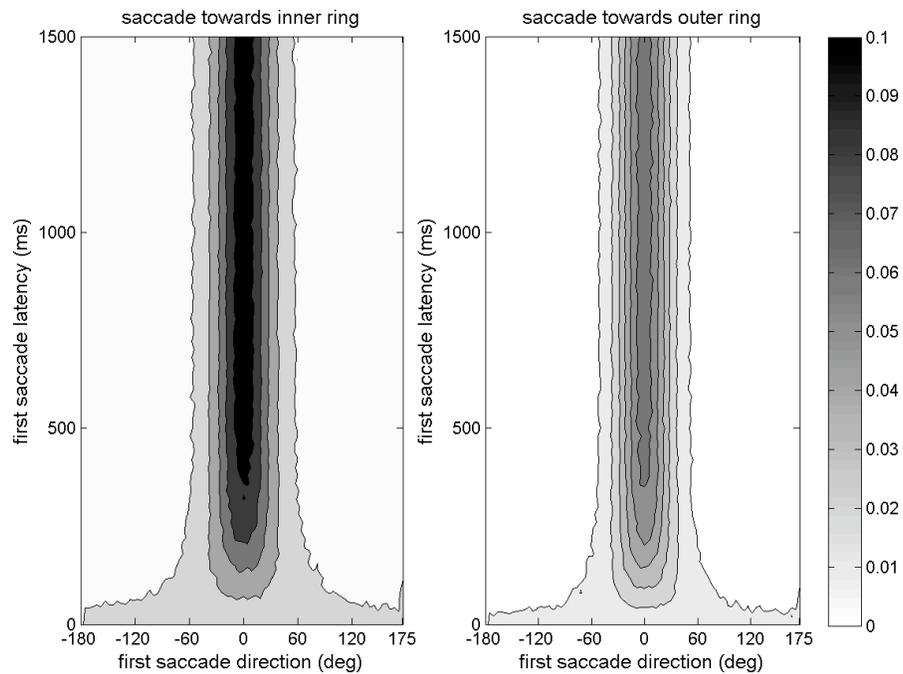


Figure 23. Contour plot of the targeting system's predicted saccade direction as a function of saccade latency and targeted ring.

Looking at the plot for saccade distance first (i.e., Figure 22), note how the mean distance pushes outward as saccade latency increases, and how it pushes outward further when the targeted object is located on the outer ring. Next notice how, even at a saccade latency of 0 ms (i.e., across the bottom of the plots), there is a non-zero average saccade distance (i.e., the plots have an intercept). Looking back at Figure 11, notice how the linear component of trend fit to the observed data predicts a non-zero average saccade distance at 0 ms latency as well (i.e., the line for both the inner and outer ring plots pass above the origin). The model accounts for this by adding a bit of noise to each object's initial activation (Equations 7 and 10). This small initial asymmetrical activation across the objects on the display causes the objects to tug at the saccade landing point a bit unevenly, which pulls the saccade slightly outward from the origin (i.e., outward from the center of mass of the display), even at a latency of 0 ms.

Looking now at the plot for saccade direction (i.e., Figure 23), notice how targeting accuracy quickly transitions from noise at onset to centering on the target (located at 0 degrees for both plots) within the first few hundred milliseconds. How quickly this transition occurs can be altered by modifying the activation rate parameter (Equation 8). Next, note how targeting accuracy is a bit better when the targeted object is on the inner ring (i.e., the width of the distribution across the x-axis is smaller for the inner ring), which occurs because the targeting system gains uncertainty about the actual location of an object as its eccentricity increases (Equations 3 and 4).

Integrating with Byrne's Saliency Model

The targeting system is in charge of moving the eyes towards the targeted object but has nothing to say about how that particular object was chosen. Byrne's visual salience model fills this gap nicely by providing the machinery that selects a salient object. After an object has been chosen, the targeting system attempts to move the eyes towards that object to minimize the time required to encode the object's attributes.

Participants certainly were not perfect in moving the first saccade towards the target. Quite often, the eyes would move towards a particularly salient distractor instead. Assuming the eyes are moving towards the attended object, imperfect eye movements imply that the visual system is imperfect in calculating each object's salience values. Byrne's salience model assumes that the object with the highest current salience (i.e., the highest L_i) is the one chosen to attend to next. Because the system is noisy when calculating these salience values, quite often a distractor object can be more salient than the target. Therefore, Byrne's salience model can potentially account for the imperfect eye movements observed.

Observed likelihoods for looking at [1] a target object, [2] a same-color distractor, and [3] a same-shape distractor were recorded for cases with two targets across six manipulated conditions (i.e., three distractor ratio conditions crossed with two ring conditions), and Byrne's salience model was fit to account for these observed frequencies. Similar to the Area Activation Model (Pomplun, Reingold, & Shen, 2003), adding a small dependence of salience on the object's eccentricity from the fovea was necessary to account for the fact that more saccades were initiated

towards objects on the inner ring overall. Formally, this scales each object's salience L_i by multiplying its value with a decaying exponential:

$$(L_i)_f = (L_i)_o \cdot e^{-qr} \quad (11)$$

Here, the i and o subscripts represent the object's initial and final salience respectively, while q is a spatial rate parameter which determines how quickly salience drops off as a function of eccentricity. The $(L_i)_o$ term in Equation 11 is simply the original L_i from Equation 1.

Visualizing the model's performance. After fitting the parameters in Equation 11 to fit the 6 manipulated conditions, the model was tested against all 18 (i.e., those 6 additionally crossed with singleton, 2 targets spaced, and 2 targets grouped). Because calculating these salience values is noisy, a Monte Carlo simulation was performed ($N = 100,000$ per cell). Also, assuming the targeting system is noisy as well – particularly with short saccade latencies (see Figures 12 and 13) – observed eye movements with only medium and long saccade latencies were used to determine the observed frequencies for cases 1-3 across each manipulated condition. The results of the fit are included in Figure 24 ($r^2(52) = .90$, $MAD = 4.2\%$).

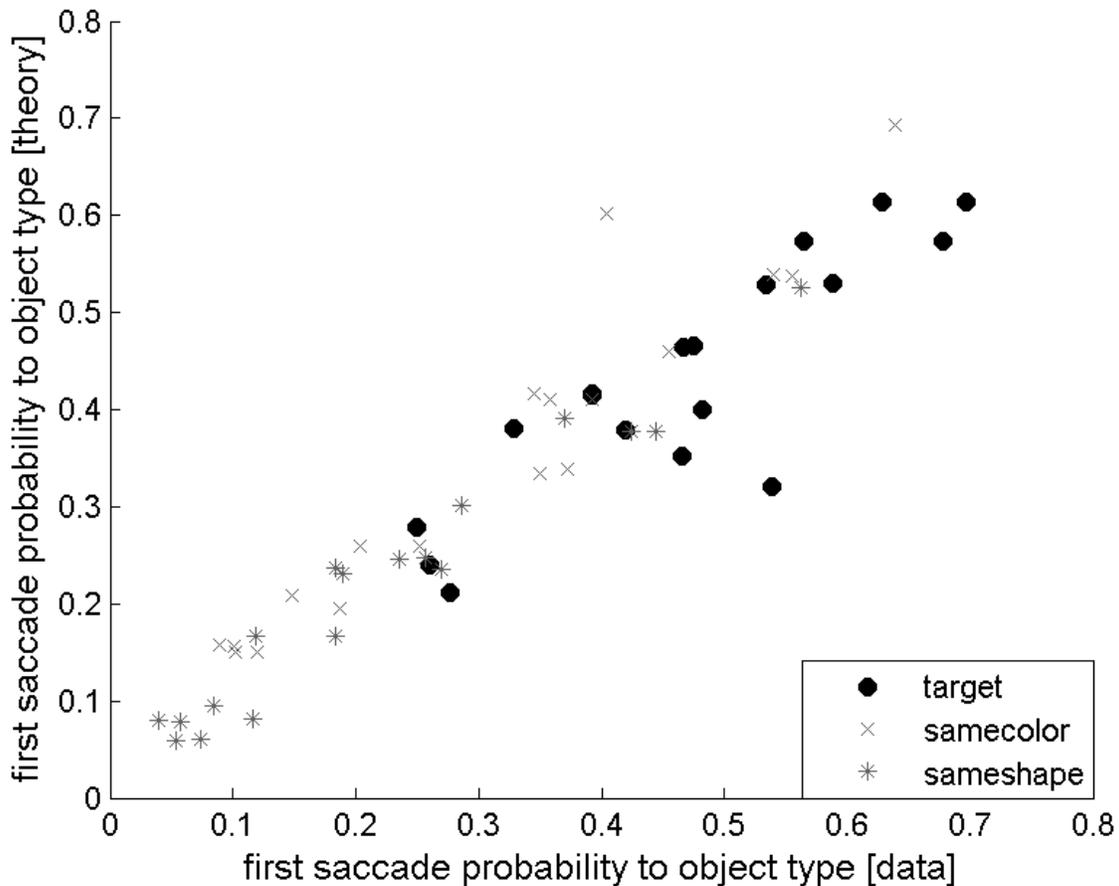


Figure 24. Scatter of predicted vs. observed frequencies for chosen first saccades.

A summary of the free parameters used for fitting these frequencies are included in Table 7 (note particularly that no bottom-up activation was required to get this high correlation).

Table 7. Free parameters used in the Byrne's salience equations to model data.

Parameter Name	Parameter Symbol	Parameter Value
Bottom-up color	γ_{color}	0
Bottom-up shape	γ_{shape}	0
Top-down color	w_{color}	1.5
Top-down shape	w_{shape}	1
Base-level sigma	s_b	.45
Transient-level sigma	s_t	.045
Eccentricity rate	q	.0415 degrees ⁻¹

Integrating with the EMMA Model

The EMMA model assumes that the landing point of a saccade is sampled from a Gaussian distribution with a mean of the object's location that the visual system is trying to encode. Assuming the participants are trying to land the saccade on the object they are encoding, a symmetrical Gaussian distribution centered on the object's location will not accurately account for the observed data here, since saccades drastically undershot targeted objects on both the inner and outer rings (see Figure 11). Therefore, the EMMA model requires a bit of modification to fit these data.

The targeting system formalized here was designed towards this end. That is, this system simply replaces EMMA's targeting component, while keeping the rest of the machinery intact. Salvucci (2001) suggests incorporating an undershoot bias into EMMA's targeting system for future revisions. An appealing property about the targeting system formalized here is that such an undershoot bias does not have to be defined explicitly; rather, the bias emerges naturally due to how the targeted and non-targeted objects collectively tug at the saccade landing point, and how the activation of the non-targeted objects takes some time to be inhibited.

The amount of time required to prepare the saccade in the EMMA model is initially set to 150 ms (with an additional 20 ms required to initiate the saccade). Looking at the distribution of observed saccade latencies (i.e., Figure 9), the mean seems to be a bit larger than what these parameters would predict. Therefore, the mean time required for the preparation stage was incremented until the predicted dispersion of latency times lined up with the distribution in Figure 9.

Visualizing the model's performance. The results of this fit are included in Figure 25 (a summary of the free parameters used are included in Table 8).

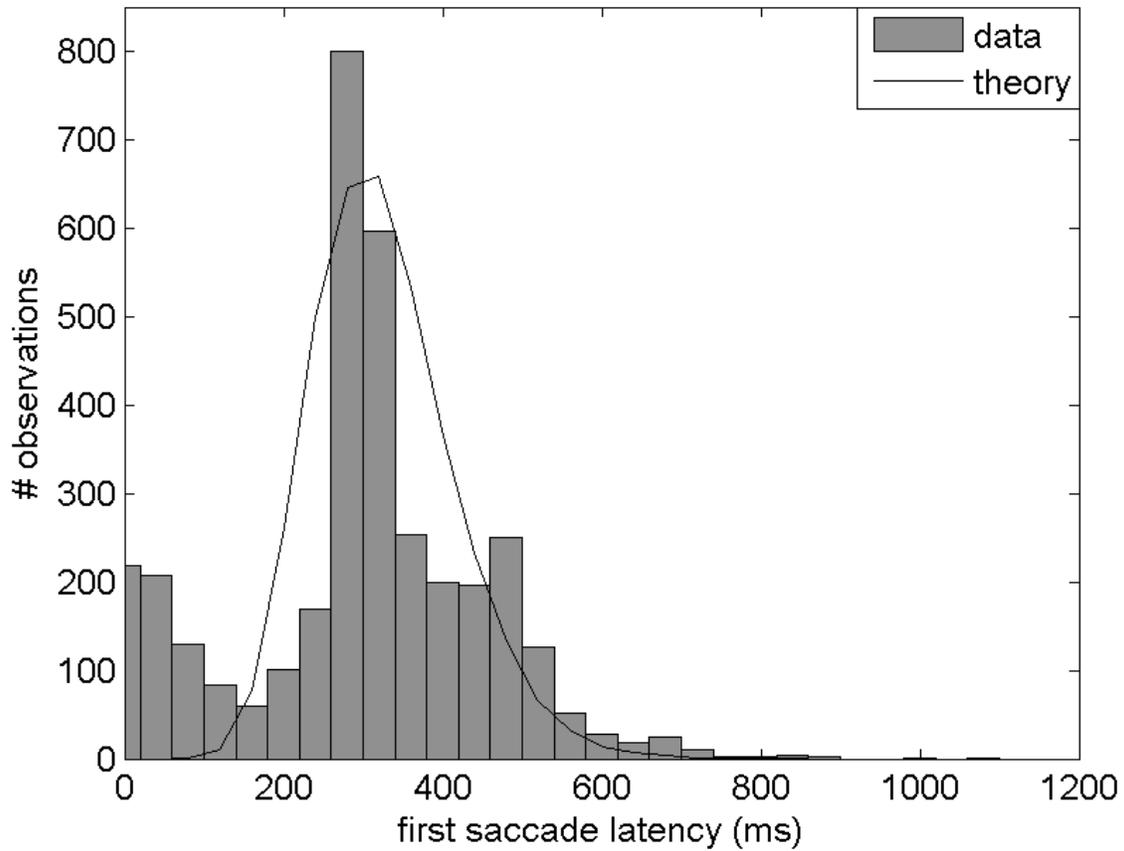


Figure 25. Fitting the observed saccade latency distribution.

Table 8. Free parameters used in EMMA to model data.

Parameter Name	Parameter Symbol	Parameter Value
Preparation mean	u_p	250 ms
Execution mean	u_e	75 ms
Preparation sigma	s_p	1/3 the mean
Execution sigma	s_e	1/3 the mean
Location sigma	s_l	1/3 the mean

Note that the two distributions in Figure 9 have been combined to form the single distribution in Figure 25 (after the saccade latency has been adjusted by 25 ms for the right distribution in Figure 9).

Running the Model

The control flow of a saccade therefore works as follows: [1]: Once the search phase begins, a request for a visual location is made. The most salient object (i.e., the object with the highest information content) is returned as the next location to attend (using Byrne's salience model). [2]: Once an object has been selected, the attributes of that object begin to be encoded; and to speed up that process, a saccade towards the object is programmed. [3]: The amount of time before the saccade is executed is sampled from the distribution in Figure 25. During this time, the targeting system is sharpening the landing point of the saccade, pushing it closer to the targeted object. [4]: Once the saccade is initiated, the landing point is calculated by taking the vector average of the targeting system's current state. If the attributes of the targeted object have not been completely encoded once this saccade ends, the process is iterated through again, pushing the eyes even closer to the targeted object, and increasing the rate of encoding accordingly.

The EMMA model describes the control flow of the next saccade once the targeted object has been encoded (several cases arising when that encoding completes during different stages). This study is only concerned with the 1st iteration of the 4 stages described, and in cases where an eye movement is executed before the object's attributes have been completely encoded. We are therefore assuming that data collected for this study had a T_{enc} greater than the latency of the first saccade, thus causing the saccade to be initiated.

A Monte Carlo simulation (N = 50,000 and 100,000 for the targeting and salience systems respectively) for each manipulated condition was performed using

the integrated model. Essentially these results were obtained by weighting the targeting system's probability distributions (one for each of the 18 objects on the 2 rings) by the likelihood that each object will be chosen for targeting. These distributions were then superimposed (keeping track of the location of each object, particularly the arrangement of target(s)), forming an *aggregated* contour plot.

Visualizing the results. Figures 26 and 27 show aggregated contour plots for saccade distance, crossed by high/low salience, inner/outer ring, and spaced/grouped target separation. Figures 28 and 29 show predicted results across the same manipulated conditions for saccade direction.

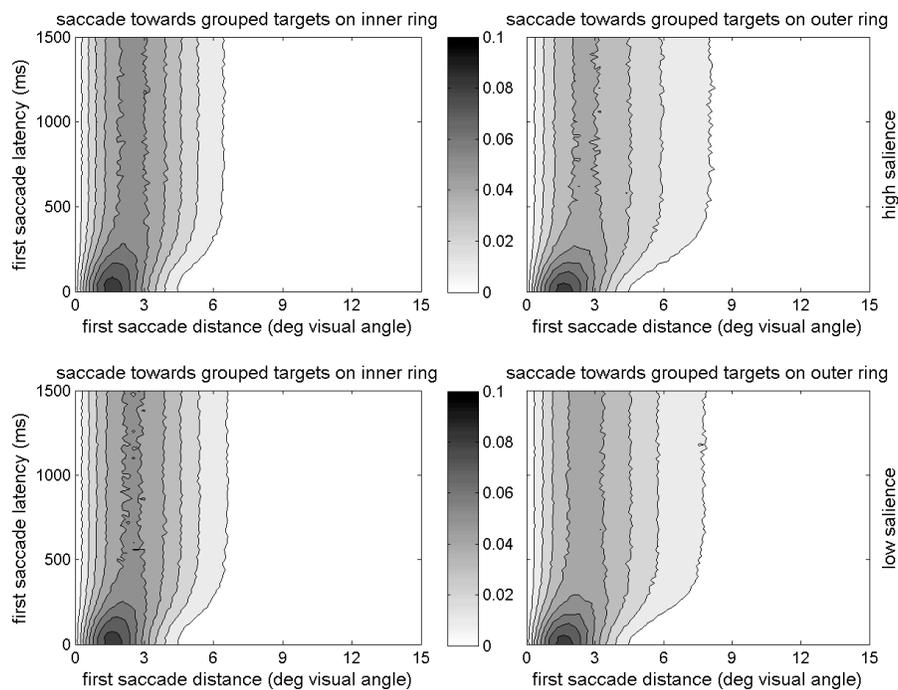


Figure 26. Contour plot of the predicted saccade distance for grouped targets as a function of saccade latency, salience, and ring.

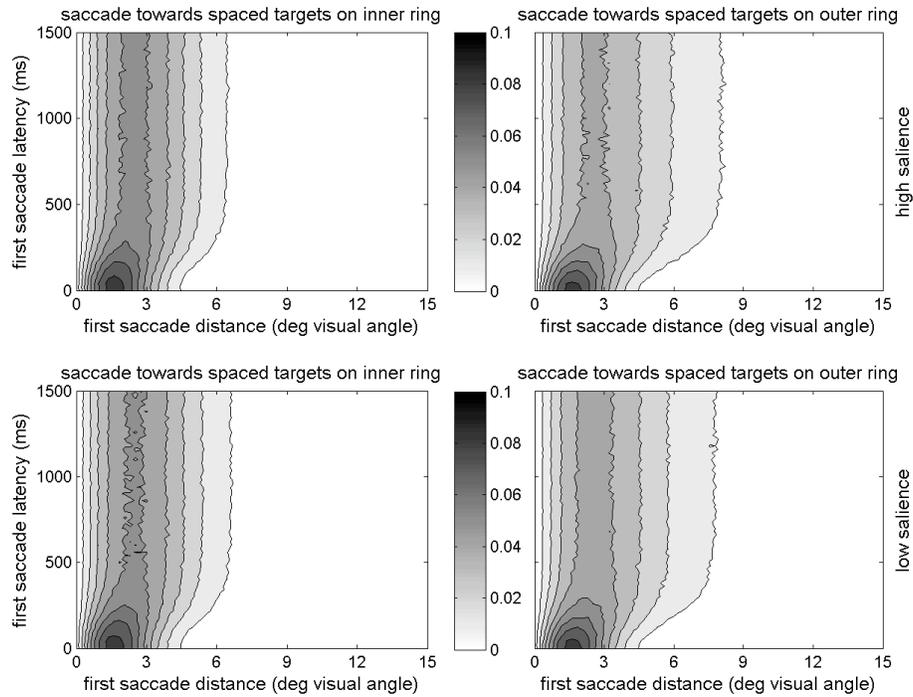


Figure 27. Contour plot of the predicted saccade distance for spaced targets as a function of saccade latency, salience, and ring.

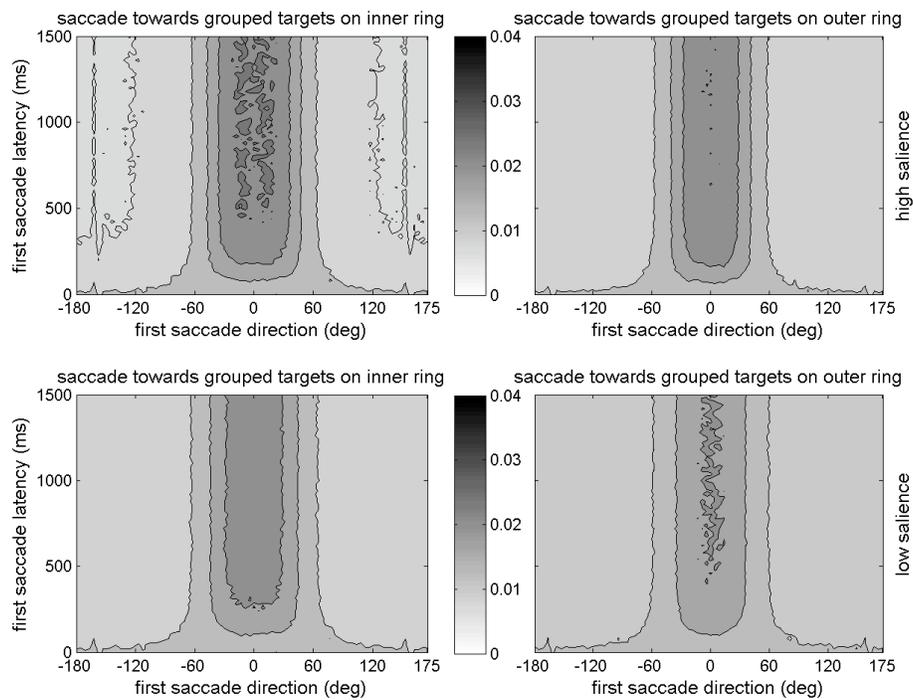


Figure 28. Contour plot of the predicted saccade direction for grouped targets as a function of saccade latency, salience, and ring.

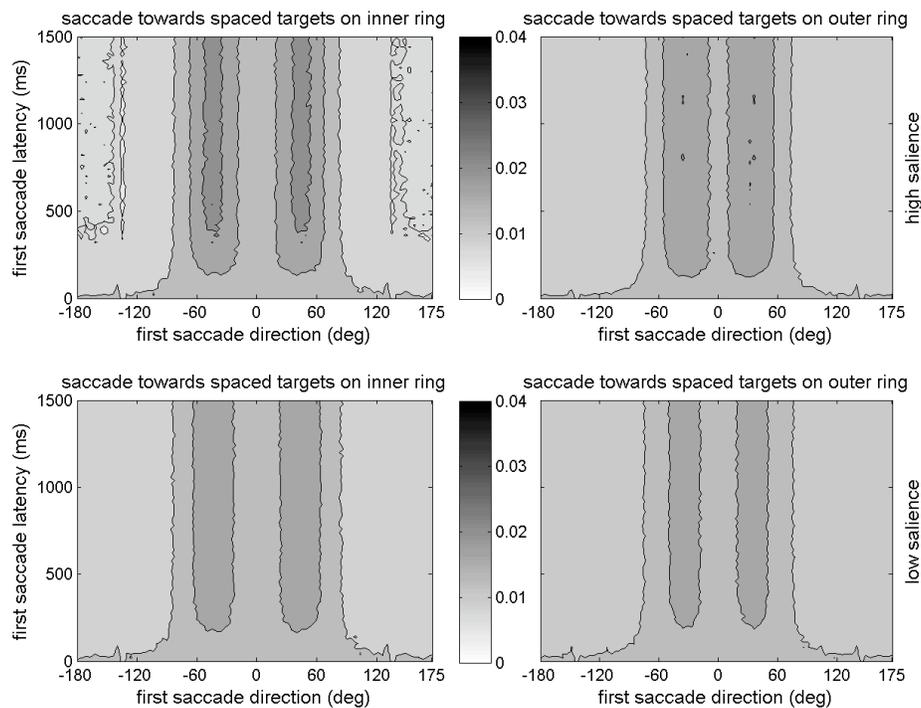


Figure 29. Contour plot of the predicted saccade direction for spaced targets as a function of saccade latency, salience, and ring.

Looking at the plots for saccade distance first, note again how the mean saccade distance pushes outward further when the target is located on the outer ring, and that a non-zero intercept is present. Also, note how the amount of the darkest region increases with increasing salience when the targets are located on the inner ring, but that this relationship flips when the targets are located on the outer ring. This is the model prediction of the observed interaction between ring and salience on saccade distance, and is due to the fact that target salience is confounded with the likelihood of looking on the inner/outer ring. For instance, when the target is on the outer ring, higher salience means that the saccade has a higher likelihood of being initiated towards an object on the outer ring, which

means that saccade distance should increase with salience. This relationship flips for targets located on the inner ring (i.e., when the target is on the inner ring, higher salience leads to more saccades to the inner ring which leads to a decrease in average saccade distance).

Next, looking at the plots for saccade direction, a much stronger effect of salience is present. Note how targeting accuracy improves as the salience of the targets increases (i.e., the darkest region is larger in high salience conditions). The relationship between targeting accuracy and eccentricity is present as well (i.e., the darkest region is larger when the targets are located on the inner ring). The large effect of latency on targeting accuracy is again present, showing a sharp transition between 0 and 200 ms. Finally, note how the distribution of saccades with relatively long latencies show 2 peaks when the targets are spaced, and 1 overlapping peak when the targets are placed side-by-side. As indicated by Figures 12 and 13, this is precisely what is observed.

Model Results

Generating model predicted means from the contour plots shown in Figures 26-29 requires collapsing across two dimensions. First, each cross section across the saccade latency dimension (e.g., all saccades at 100 ms) is itself a probability density distribution, having a normalized area of 1. The mean of this cross section is therefore the sum of all values on the cross section weighted by each value's corresponding probability (i.e., the dot product between the x-axis and the height of the contour). After performing this task across each saccade latency cross-section, a

function relating the average value of the given variable (say saccade distance) and latency arises.

However, it is not the case that all saccade latencies are equally likely (see Figure 25). Therefore, this function also has to be weighted by each value's corresponding probability (i.e., the likelihood of having that latency) when determining the mean (i.e., the dot product between this function and probability distribution in Figure 25 is taken). The mean predicted by the model for a particular variable is therefore the result after collapsing across both these dimensions (observed value and latency).

The dependent variables chosen for testing were saccade distance and saccade directional error. Although the directional error contour plots were not shown here, calculating the predicted means was done in the same collapsing manner as previously discussed.

Free parameters in the model were manipulated to fit 4 observed means for each dependent variable (i.e., low/high salience crossed with inner/outer ring). However after fitting, model predictions were tested against all 18 observed means for each dependent variable (i.e., crossing all 3 distractor ratio conditions and all 6 relative spacing conditions).

Effects on Saccade Distance

Effect of ring. A large effect of ring on saccade distance is present, and the model correctly predicts the direction and magnitude (roughly) of the observed means (Figure 30).

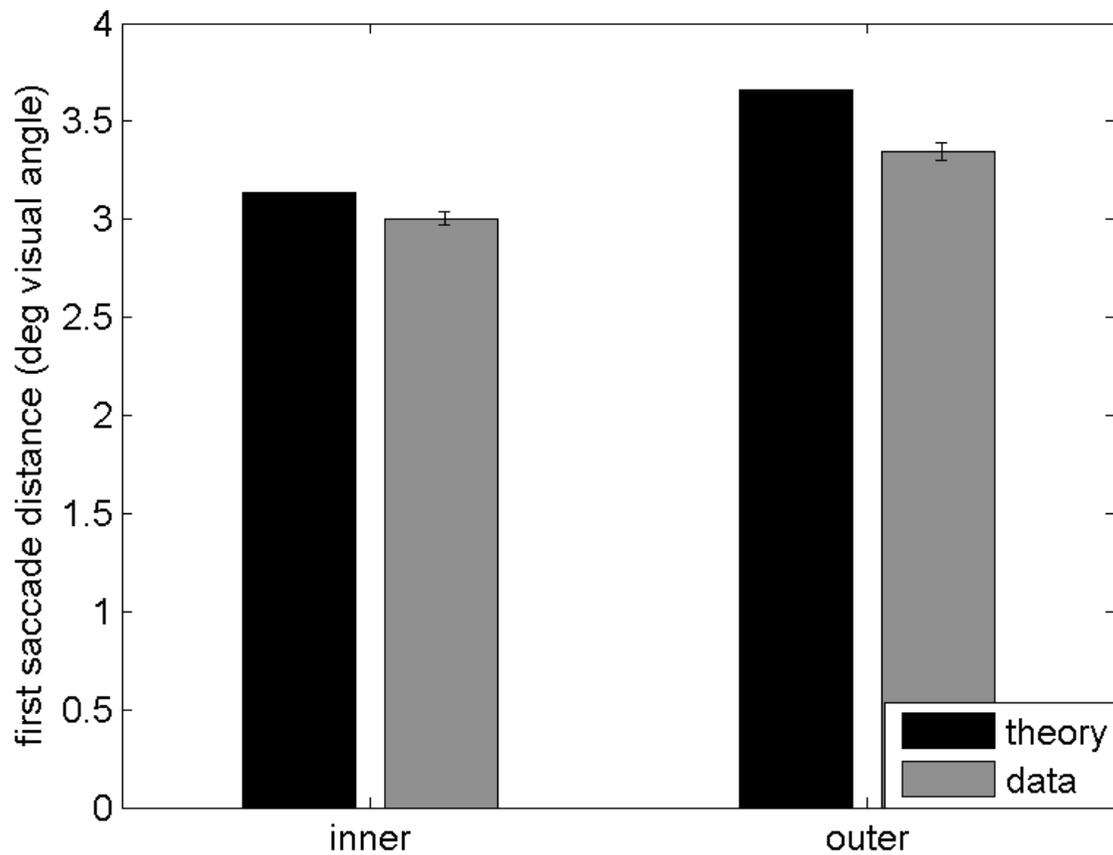


Figure 30. Predicted and observed means for the effect of ring on saccade distance.

Effect of salience. A null effect of salience on saccade distance is present, and the model correctly predicts having roughly no change across this dimension, although the intercept is slightly high (Figure 31).

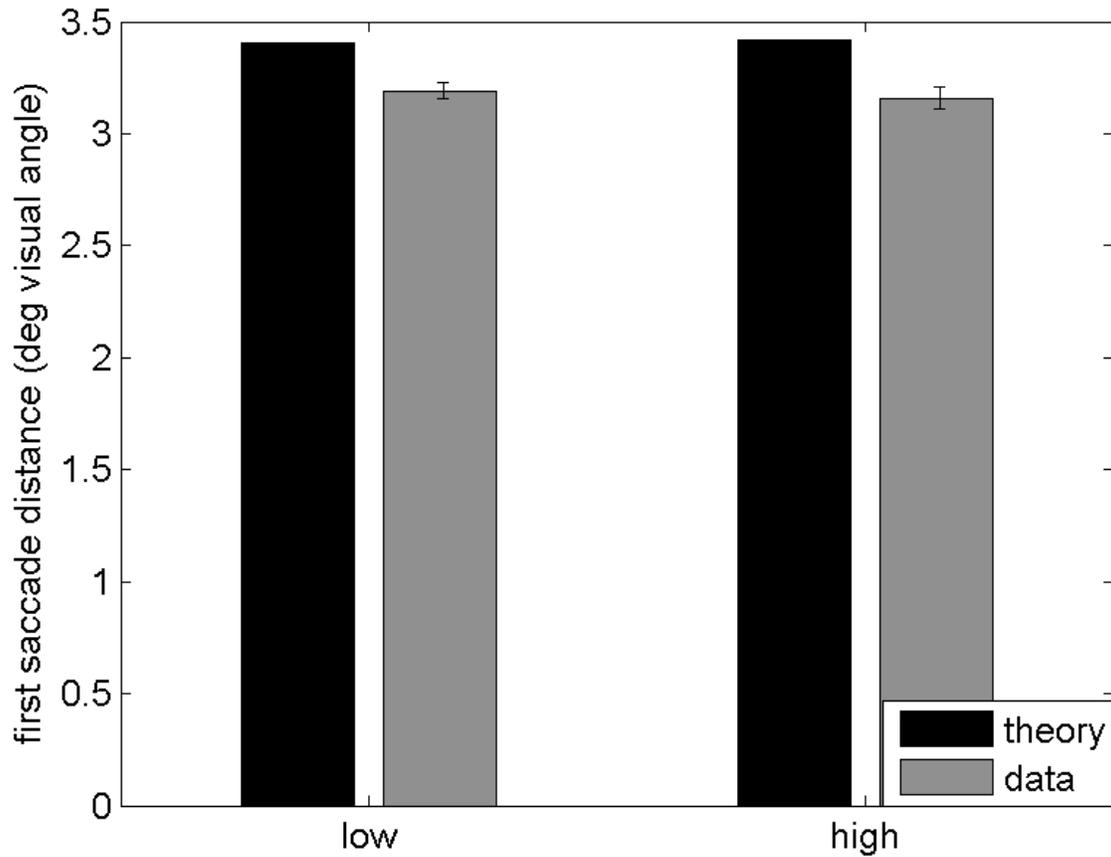


Figure 31. Predicted and observed means for the null effect of salience on saccade distance.

Interaction of salience and ring. The model correctly predicts the directional change for the effect of salience on saccade distance dependent upon the ring condition, although again the intercept is high (Figure 32).

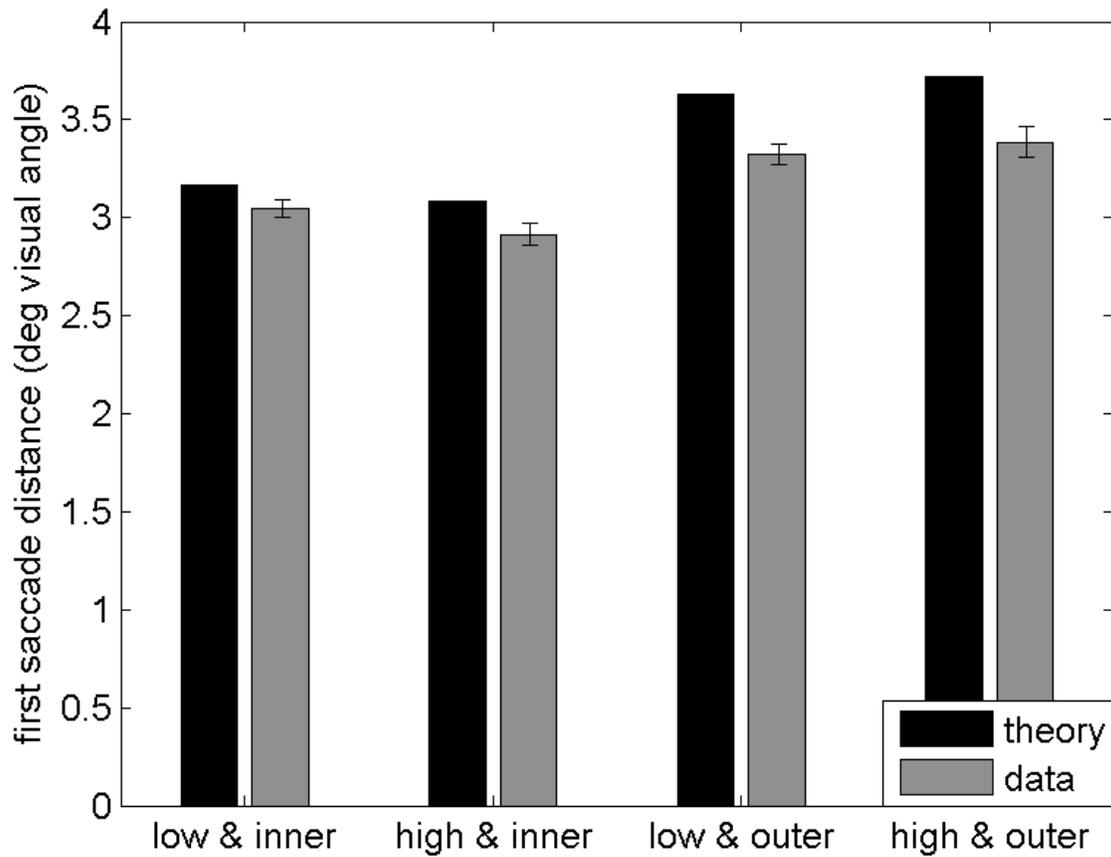


Figure 32. Predicted and observed means for the interaction between salience and ring on saccade distance.

Scatter. A large range effect is present for the effect of saccade distance on ring, so it becomes increasingly difficult to tease apart the effects of other experimental manipulations (Figure 33, $r^2(16) = .83$, MAD = .23 degrees).

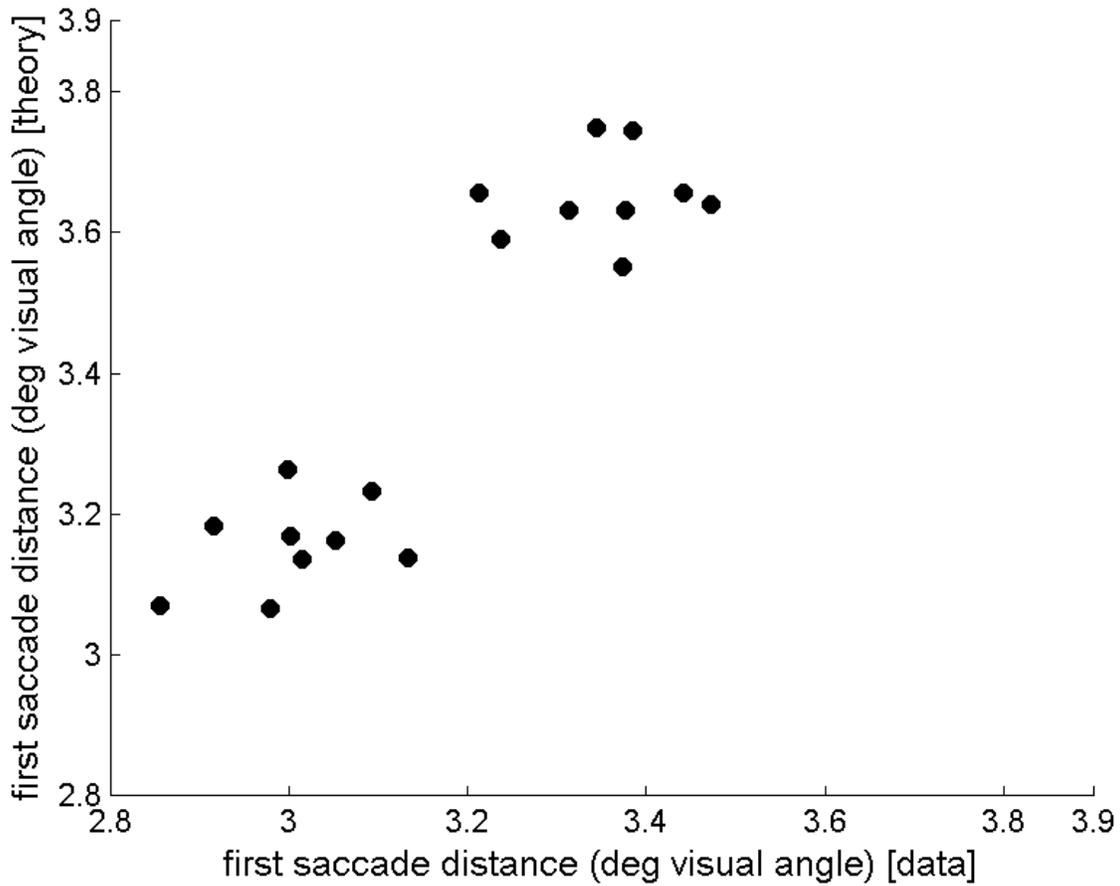


Figure 33. Scatter of predicted vs. observed means for saccade distance.

Effects on Saccade Directional Error

Effect of ring. A large effect of ring on saccade directional error is present, and the model correctly predicts the direction of the observed means, although the model has slightly less targeting accuracy for each mean (Figure 34).

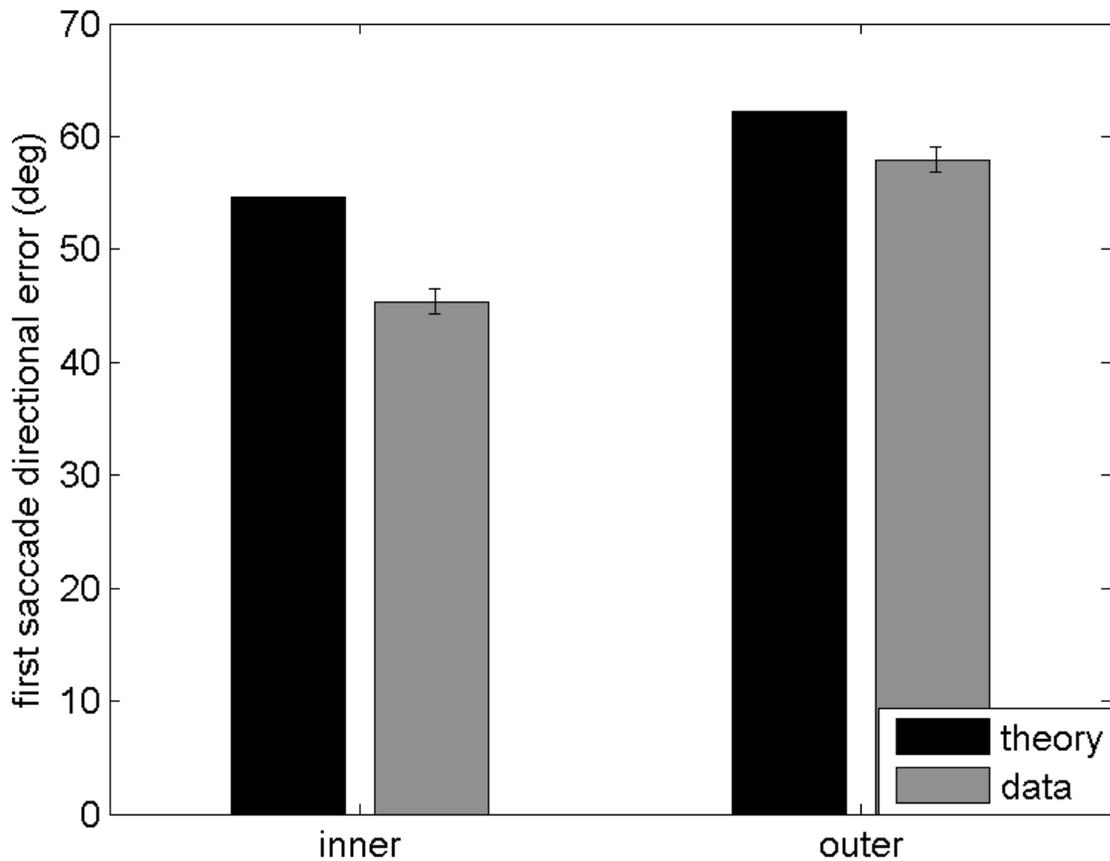


Figure 34. Predicted and observed means for the effect of ring on saccade directional error.

Effect of salience. Targeting accuracy should improve when the target is more salient, and the model correctly predicts this trend (Figure 35).

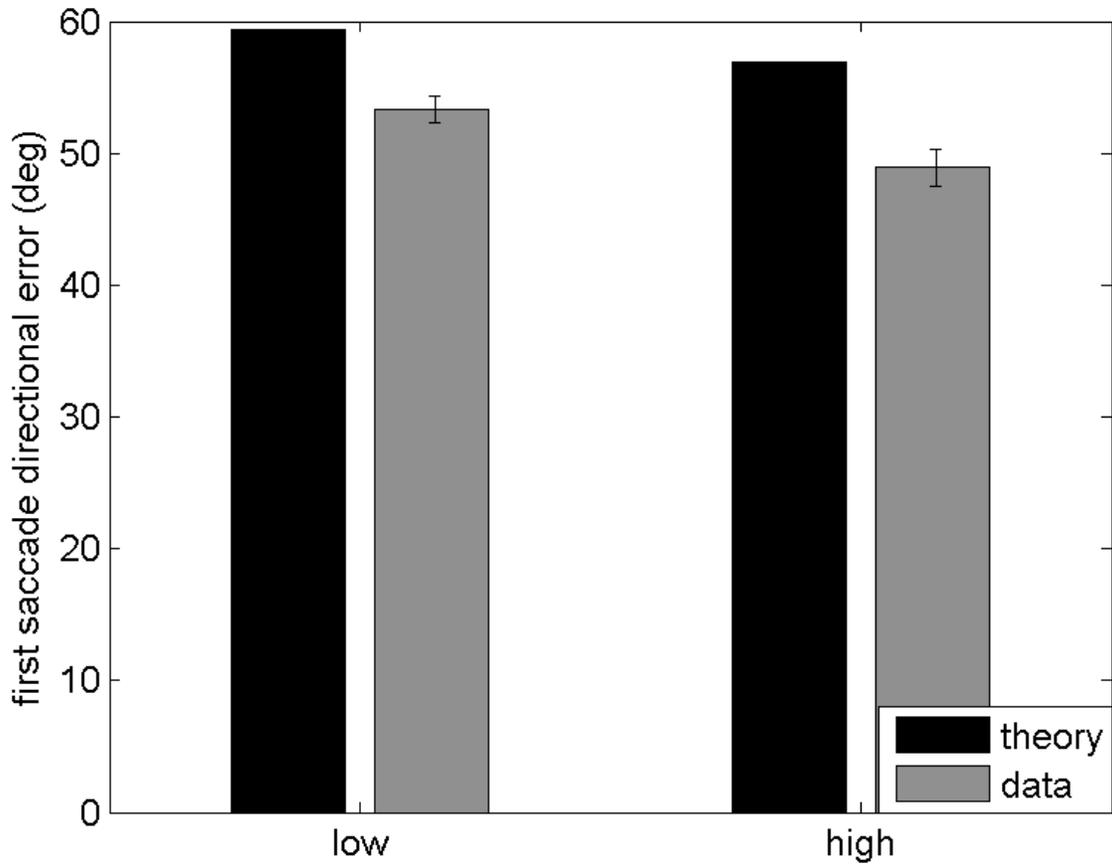


Figure 35. Predicted and observed means for the effect of salience on saccade directional error.

Interaction of salience and ring. Salience effects targeting accuracy more when the target is on the inner ring, and it is questionable whether the model correctly predicts this interaction (Figure 36).

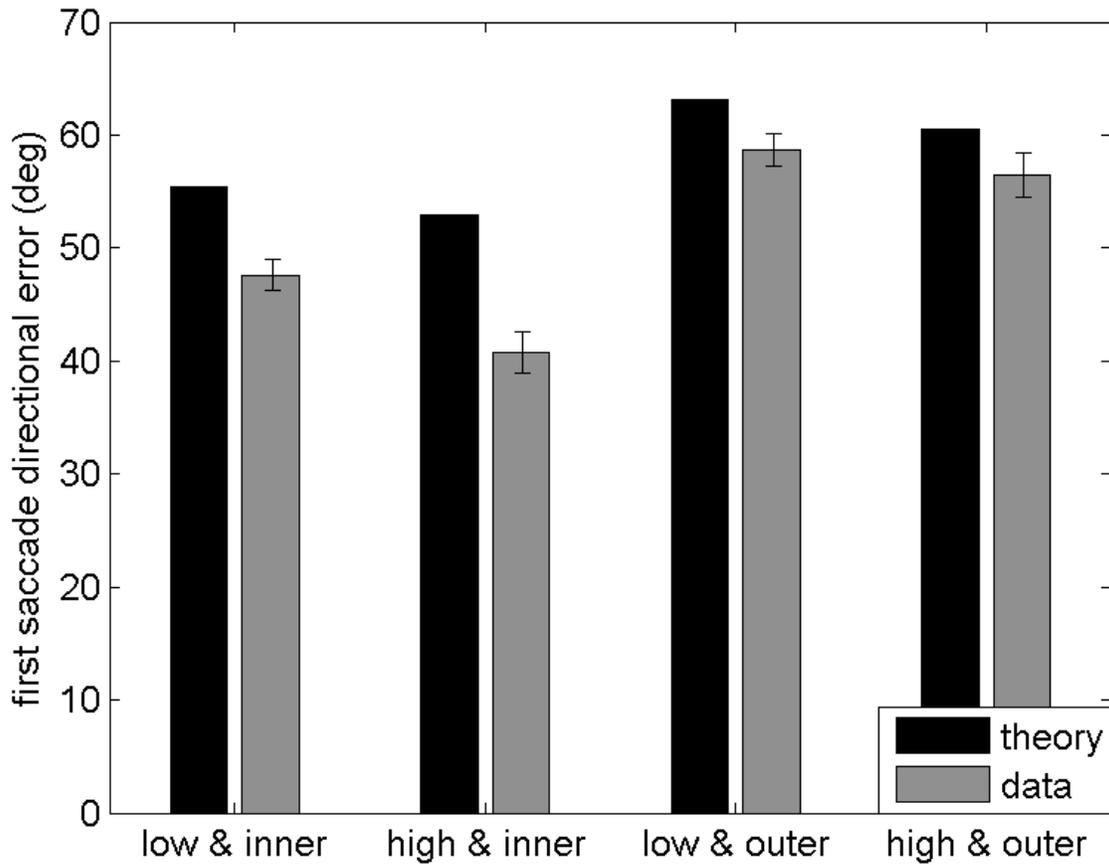


Figure 36. Predicted and observed means for the interaction between salience and ring on saccade directional error.

Scatter. The means for saccade directional error are more evenly dispersed than the means for saccade distance (Figure 37). Variance predicted by the model in this case was high, even without the help of a large range effect ($r^2(16) = .88$, MAD = 7.68 degrees). However, it is apparent in the scatter that the model has a bit less targeting accuracy overall (i.e., a higher intercept).

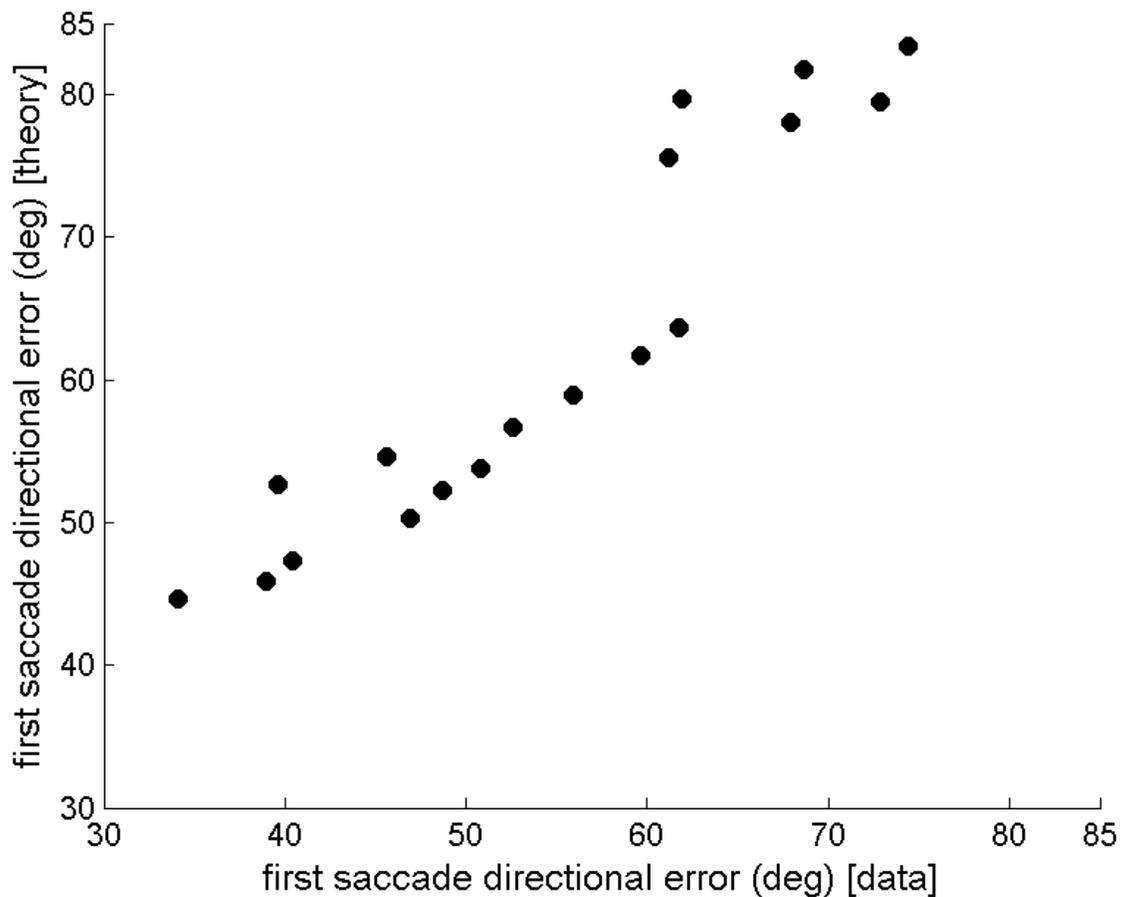


Figure 37. Scatter of predicted vs. observed means for saccade directional error.

Saccades with Questionably Short Latencies

Looking back at the observed distribution of saccade latencies (Figure 9 as well as Figure 25), note the small clustering of relatively quick saccades, arguably forming a 2nd peak in the distributions that is pushed against the origin. The EMMA

model outlines a 2-stage system to make a saccade (i.e., a preparation and execution phase). If a saccade has been programmed, and the visual system wishes to cancel the saccade and move the eyes to a different location, this cancellation is only possible if the visual system is still in the preparation phase. If it is instead in the execution phase, the initially programmed saccade cannot be cancelled, and the visual system will therefore have to wait until that eye movement is finished before requesting to move to the new location. The EMMA model sets this execution phase to lasting ~ 50 ms. Therefore, one could argue that saccades made in the first 50-75 ms after the search trial begins are most likely saccades corresponding to programming *before* the onset of the search phase that could not be canceled when the search phase began. These saccades should therefore not be included, since their guidance was not determined by the configuration of the search phase.

Towards this end, a 2nd run of the model was performed, where any saccade less than 100 ms was trimmed from the dataset (~ 500 of the 3500 observations were excluded). All free parameters remained the same, and the results did slightly sharpen qualitatively (i.e., the range effect of ring on saccade distance was notably smaller, and the slope was much closer to unity; compare Figure 33 to Figure 38), although the correlations and absolute deviations did not change much (see Table 9 for a side-by-side comparison), and got slightly worse in some cases.

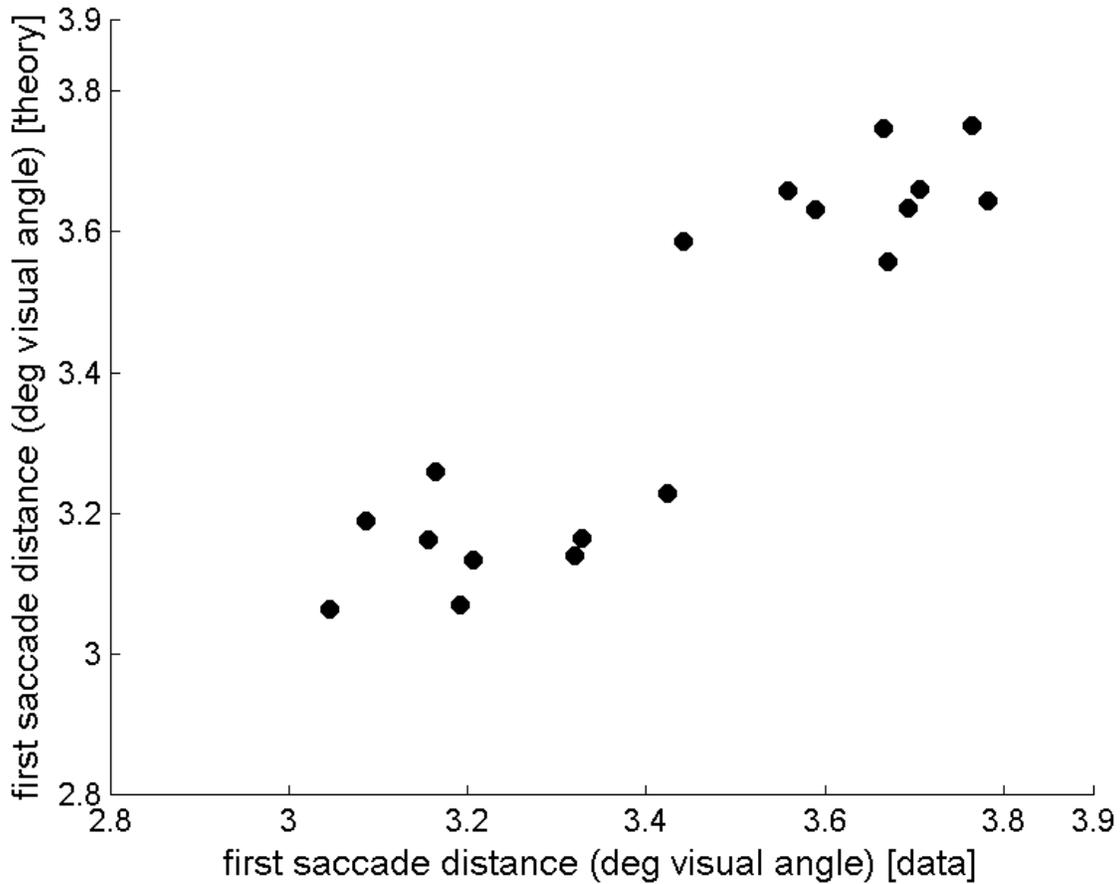


Figure 38. Scatter of predicted vs. observed means for saccade distance after trimming observations with <100 ms saccade latencies.

Table 9. Model performance with and without trimming < 100 ms saccade latencies.

	Entire dataset	<100 ms removed
Targeted frequencies $r(52)$.95	.95
Targeted frequencies MAD	.042	.043
Mean saccade distance $r(16)$.91	.91
Mean saccade distance MAD	.23 degrees	.10 degrees
Mean saccade directional error $r(16)$.94	.92
Mean saccade directional error MAD	7.68 degrees	10.20 degrees

A large portion of the dataset was trimmed during this process (roughly 1/7).

Combined with the fact that the results were not much improved, trimming these data becomes increasingly difficult to justify. Also it is unclear if these short-latency

saccades are truly due to this cancellation failure described, for the participants were not sporadically moving their eyes across random parts of the screen before each search trial began. Rather, most subjects kept their eyes centered on the target to find (i.e., centered on the middle of the display) until the search phase began.

DISCUSSION

Interpretation of the “global effect averaging” and “center of gravity” phenomena

Speed/Accuracy tradeoff. Saccade latency strongly influences the accuracy of the saccade to a targeted object. Both the observed results (see Figures 12 & 13) and the model (see Figures 28 & 29) show a clear transition from noise to accurate targeting as saccade latency increases.

Findlay (1997), using an analogous concentric ring experiment, found evidence for “global effect averaging” (i.e., eye movements landing in between two target objects) only with saccades less than ~ 300 ms. Looking back at Figures 12 and 13, this experiment had slightly noisier results. This was most likely due to having two concentric rings for this experiment, causing a bit of uncertainty regarding which ring the saccade was initiated towards. Figures 12 and 13 were generated assuming the saccade was made towards the ring that the target was located on. However, Byrne’s salience equations tell us that this certainly is not always the case, meaning that a portion of the noise for each plot in these figures is due to guessing the wrong ring that the saccade was moving to.

Nevertheless, the general trend of accuracy increasing with latency is present here as well, and the model provides insight as to why this is occurring. Specifically, because the activation of non-targeted objects takes a bit of time to be inhibited, and

the activation of the targeted object takes a bit of time to fully saturate, saccades initiated with short latencies have not given the targeting system enough time to develop a high signal to noise ratio. Consequently, non-targeted objects still carry a large amount of weight, pulling hard on the saccade to move to their location, which in turn makes the landing point of the saccade much noisier (and closer to the origin on average).

The task studied by Findlay (1997) was a symmetrical arrangement of objects with the initial fovea location already at the objects' center of gravity. In this case, any activation from non-targeted objects works against the saccade spanning outward towards a targeted object and thus creates an undershoot bias. However, in a symmetrical arrangement of objects with the initial fovea location *not* at the objects' center of gravity (e.g., Rao et al., 2002), any activation from non-targeted objects works to drive the saccade towards the objects' center of the gravity instead of towards a targeted object. Therefore, the targeting model here would predict the first saccade in Rao's study to move towards the center of the display, and each successive saccade to hone closer on the targeted object, which was precisely what Rao et al. (2002) observed.

Therefore, it seems that this "center of gravity" phenomenon (observed in Rao et al.'s study) and "global effect averaging" phenomenon (observed in Findlay's study) are caused by the same time-evolved system, calculating the saccade trajectory by taking the vector average of some targeting activation map. Rao et al. argued that this activation map is closely linked to the salience of the objects displayed. In the case here, this targeting map is completely decoupled from the

saliency map. Nevertheless, in both cases a speed/accuracy tradeoff arises because of this system, whereby saccades with short latencies are much noisier than saccades given a bit more time to program.

Thus, these observed phenomena are not tied to a specific task paradigm, nor a particular display configuration. Rather, this tendency to move the eyes towards the center of mass of various objects on the display, both locally (i.e., between 2 grouped objects) and globally (i.e., between all the displayed objects), arises due to the same underlying machinery causing a speed/accuracy tradeoff to exist, sharpening saccade accuracy as latency increases.

Future Directions

Distinguishing between intense and distinctive items. Byrne's theory of visual saliency computation has only been tested on visual displays containing objects with similar perceptual properties (e.g., similar brightness) and varying distinctive properties (e.g., shape, orientation, color). It may certainly be the case that an object's perceptual properties interact with its distinctive properties to determine its information value. Therefore, it would be useful to examine a visual search task where target and distractor arrangements are manipulated across both of these dimensions, so that the contribution of basic perceptual properties on an object's information value can be included in the visual saliency theory as well. For example, one could make the target the dimmest object, or train participants to find the object that has the *least* information content. In a sense, these types of tasks aim to address how much conscious or strategic control an individual has when searching for an

object (i.e., top-down activation) compared to the amount that attention is driven by basic perceptual properties alone (i.e., bottom-up activation).

Fitting parameters to match RTs. Parameters here were fit using observed data for the first saccade only. However, if the model is run through the entirety of the search task (i.e., through all the saccades generated and cognitive cycles), the reaction times predicted before concluding present/absent should line up with the reaction times observed. In other words, using the same set of parameters, the model should not only be able to predict the observed frequencies of the first saccade looking at targets/distractors, but also the total amount of time spent on the task before a present/absent conclusion is reached. This further constrains the space of possible parameters, for the model now has to be fit to a larger collection of observed variables.

Of particular interest is the correct weighting of bottom-up and top-down activation parameters when RTs must be matched alongside first saccades. Only top-down activation was used to fit data for this task. However, bottom-up activation was necessary to fit a similar task studied by Stanley and Byrne (2008). Further work is needed to resolve this dissonance in parameter values, and hopefully converge on a given set of values that capture the majority of the variance observed across both eye-movement data and RTs for a reasonable range of tasks.

Building an ACT-R implementation. A mathematical model in Matlab was used to generate the predicted results for this study. However, if one wishes to test the model performing the entirety of the search task (and not just the first saccade),

implementing the theory into a pre-existing cognitive architecture and building a model of the task is the most reasonable next step.

ACT-R seems the most appropriate, since both the EMMA model and Byrne's salience model already have an ACT-R implementation. All that is really needed is an additional module tying those theories together, as well as replacing EMMA's targeting system with the one formalized here, which would take less than a few hundred lines of code to do.

Conclusions

It is important to realize that the computational model formalized here is built from a relatively simple set of sequential and approximately independent systems: a salience map determines where visual attention should move, which provides input to a saccadic system that controls the programming, executing, and canceling of saccades, which uses a targeting system to calculate the saccade's polar distance and direction. However, when these relatively simple systems are integrated and examined together, the computational model has a large amount of explanatory power, providing insight concerning how saccade latency, age of objects, signal to noise ratio, and spatial arrangements of targets and distractors all interact to influence saccade landing points and response times in visual search tasks. The model is also computationally efficient (i.e., Byrne's salience model, EMMA, and the targeting model all operate on the order of the number of *objects* displayed, and not on the number of *pixels* displayed), and can therefore run at or faster than real time. However, the model still predicts effects at a high resolution (i.e., the landing point of saccades at the *pixel* level). Finally, because each

component is approximately independent, one can switch out or supplement various pieces with other computational models if one desires.

For example, Wolfe's Guided Search 4.0 (2007) or Itti and Koch's saliency-based search mechanism (2000) could replace Byrne's salience model in calculating each object's salience. However, this was not pursued here since Byrne's salience model already predicted observed targeted objects to a high level of accuracy across all experimental conditions ($r^2(52) = .90$, $MAD = 4.2\%$; see Figure 24).

Alternatively, if one has good reason to argue that the targeting system is salience-sensitive, Logan's CODE theory of visual attention (1996), using $v(x,i)$ as the rate parameters in the targeting activation network, could enable this behavior in the targeting model here. Using this method, each object's targeting activation would again start at the same initial level. However, instead of having the targeted/non-targeted objects increase/decrease in activation with increased latency, every object's activation would increase with latency but at a rate proportional to its salience. Formally, each object's targeting rate parameter $v(x,i)$ and salience would be linked using the following relationship:

$$k_i = v(i,T) = \eta(i,T)\beta_T \frac{w_i}{\sum w_z} \propto [(L_i)_f]^2 \quad (12)$$

where T stands for target, k_i is the rate parameter in Equation 8 for object i , and $(L_i)_f$ is the salience of object i calculated by Equation 11. However, this was not implemented primarily because saccade distance was not largely influenced by the relative salience of the target object ($F(1,15.06) = .90$, $p = .36$; see Figure 14), and Equation 12 would predict a strong effect of salience on saccade distance.

The general purpose of this study was to increase understanding of the underlying processes causing the global effect. By finding that task presentation only minimally influences targeting accuracy, one can argue that the majority of the temporal dynamics causing the global effect are not located in the system which determines the salience of the objects. By finding that target salience largely influences targeting accuracy but not targeting distance, one can argue that these temporal dynamics instead reside in a separate targeting system that is not largely concerned about what the object is, and instead cares only about where the objects are and which object it's been told to look at. By equating this global effect to a simple speed/accuracy tradeoff, one gains a clearer understanding as to why the global effect occurs, and how it arises in both local and global forms. Finally, by realizing that the global effect is largely dependent on the latency of the saccade, which is *not* largely dependent on the salience of objects, the results pose future questions concerning how this saccade latency can be manipulated to potentially influence eye movement accuracy.

REFERENCES

- Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research, 14*, 589-592.
- Bacon, W. F., & Egeth, H. E. (1997). Goal-directed guidance of attention: evidence from conjunctive visual search. *Journal of Experimental Psychology, 23*, 948-961.
- Banks, B. D., Mao, I. L., & Walter, J. P. (1985). Robustness of the restricted maximum likelihood estimator derived under normality as applied to data with skewed distributions. *Journal of Dairy Science, 68*, 1785-1792.
- Byrne, M. (2006). Rational analysis of visual salience and search. Presented at the meeting of the Biologically-Inspired Cognitive Architectures (BICA) Program. Retrieved August 6, 2008, from <http://chil.rice.edu/projects/salience/>
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, 39*, 1-38.
- Findlay, J. M. (1997). Saccade target selection during visual search. *Vision Research, 37*, 617-631.
- Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: the psychology of looking and seeing*. New York, NY: Oxford University Press.
- Hirsch, J., & Curcio, C. A. (1989). The spatial resolution capacity of human foveal retina. *Vision Research, 29*, 1095-1101.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 11-46.

- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology, 4*, 219-227.
- Logan, G. D. (1996). The CODE theory of visual attention: an integration of space-based and object-based attention. *Psychological Review, 103*, 603-649.
- Ottes, F. P., Van Gisbergen, J. A. M., & Eggermont, J. J. (1985). Latency dependence of colour-based target vs nontarget information by the saccadic system. *Vision Research, 25*, 849-862.
- Pomplun, M., Reingold, E. M., & Shen, J. (2003). Area activation: a computational model of saccadic selectivity in visual search. *Cognitive Science, 27*, 299-312.
- Rao, R., Zelinsky, G., Hayhoe, M., & Ballard, D. (2002). Eye movements in iconic visual search. *Vision Research, 42*, 1447-1463.
- Rayner, K., & Morrison, R. E. (1981). Eye movements and identifying words in parafoveal vision. *Bulletin of the Psychonomic Society, 17*, 135-138.
- Salvucci, D. D. (2001). An integrated model of eye movements and visual encoding. *Cognitive Systems Research, 1*(4), 201-220.
- Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: word frequency effects and individual differences. *Memory and Cognition, 26*, 1270-1281.
- Shen, J., Reingold, E. M., & Pomplun, M. (2003). Guidance of eye movements during conjunctive visual search: the distractor-ratio effect. *Canadian Journal of Experimental Psychology, 57*, 76-96.
- Stanley, C., & Byrne, M. (2008). Processes influencing visual search efficiency in

- conjunctive search: a rational analysis approach. *Fifteenth Annual ACT-R Workshop Proceedings*. Retrieved August 6, 2008, from <http://chil.rice.edu/projects/salience/>
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97-136.
- Wolfe, J. M. (1994). Guided Search 2.0: a revised model of visual search. *Psychonomic Bulletin & Review, 1*, 202-238.
- Wolfe, J. M. (2007). Guided Search 4.0: current progress with a model of visual search. In W. Gray (Ed.), *Integrated Models of Cognitive Systems* (pp. 99-119). New York, NY: Oxford University Press.