Spring 2015

Learning to overcome distraction

Daniel Brown Vatterott
University of Iowa

LEARNING TO OVERCOME DISTRACTION

by

Daniel Brown Vatterott

A thesis submitted in partial fulfillment of the
requirements for the Doctor of
Philosophy degree in Psychology
in the Graduate College of
The University of Iowa

May 2015

Thesis Supervisor: Professor Shaun P. Vecera

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Daniel Brown Vatterott

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Psychology at the May 2015 graduation.

Thesis Committee:

Shaun P. Vecera, Thesis Supervisor

Cathleen Moore

J. Toby Mordkoff

Andrew Hollingworth

Daniel Tranel

ACKNOWLEDGEMENTS

I would like to thank my advisor, Shaun Vecera, for providing a supportive, productive environment. Shaun provided the support I needed to push forward and the guidance I needed to know where to push. These five years have been a great ride and they would not have been the same without such a supportive mentor. I would like to thank the lab: Zach Roper, Will Bush, Hannah Wyland, and Josh Cosman. We have had some great discussions. From attention to Bitcoins, Morels, and virtual reality, I have enjoyed every moment. I would like to thank my committee: Cathleen Moore, Toby Mordkoff, Andrew Hollingworth, and Dan Tranel. Not only did all of you help improve the work presented here, but you also helped shape my thinking about cognitive psychology. I would like to thank Becky Huber for her tireless dedication to this department.

Finally, I would like to thank my family, especially my parents who have continually supported me throughout my entire education: a gift that I value above all others. Lastly, I would like to thank my fiancé, Amy, for hearing out my endless complaints and providing the critical input needed to see this work through.

ABSTRACT

Complex behaviors require selectively attending to task-relevant items, and ignoring conspicuous, irrelevant items. For example, driving requires selectively attending to other cars on the road while ignoring flashing billboards. Dominant models of attentional control posit that we avoid distraction by biasing attention towards task-relevant items, and our ability to avoid distraction depends on the strength and specificity of this bias. I find that a strong, specific bias towards task-relevant items is insufficient for preventing distraction. Instead, preventing distraction also requires past experience ignoring distractors. I also find that long-term memory systems, rather than visual short-term memory or priming memory systems, maintain this experience. Based upon these findings, I propose that effective attentional control not only demands a strong, specific bias towards task-relevant items, but also requires that observers *learn* to ignore conspicuous, irrelevant items.

PUBLIC ABSTRACT

When driving, maintaining attention on task-relevant items, such as other cars on the road, is critical for avoiding dangerous events. Attention researchers posit that maintaining attention on task-relevant items depends solely on our ability to discriminate between relevant and irrelevant items. For instance, when we can easily discriminate between relevant items, such as cars, and irrelevant items, such as billboards, we selectively attend the relevant items. When we cannot discriminate between relevant and irrelevant items, we attend to both the relevant and irrelevant items. Here, I test whether target discriminability is sufficient for preventing distraction of if we also must learn to ignore conspicuous items. I find that even when participants can easily discriminate between relevant and irrelevant items, participants must learn to ignore conspicuous items. Without this learning, the conspicuous items distract participants. I call this learned distractor rejection. I go on to investigate how participants maintain learned distractor rejection. I find that the learned ability to ignore conspicuous distractors persists over extended delays, demonstrating that long-term memory systems support learned distractor rejection. These results indicate that our ability to maintain attention on task-relevant items not only depends on our ability to discriminate between relevant and irrelevant items, but also our learned ability to ignore conspicuous items and this learning persists over extended periods of time.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Cherry (1953) found that when observers repeated an audio stream played to their right ear, they could not report the language of a simultaneous audio stream played to their left ear. This inability to report even the language of the audio stream played to observers' left ear demonstrates that our sensory systems cannot fully process all the information they receive (i.e., limited processing capacity). Yet, observers' ability to accurately repeat the audio stream played to their right ear demonstrates our ability to selectively process a portion of this information (i.e., selection). This famous experiment exemplifies our limited processing capacity and our ability to select some information to the exclusion of other information. Cognitive researchers named this *attention*. In this project, I will define attention as *the selective processing of information.* This project examines how the visual system controls were attention is allocated (i.e., attentional control). In other words, how does the visual system choose what to select and what to ignore?

Researchers often describe attentional control as a decision of where to direct attention. This leads to the important question - what information guides attention? S*timulus-based factors* (i.e., bottom-up selection) and *goal-driven factors* (i.e., top-down selection) guide attention (Folk et al., 1992; Theeuwes, 1992). Stimulus-based factors such as the large color difference between red and green, allow observers to immediately locate a red circle among green circles (Treisman & Gelade, 1980). Goal-driven factors such as task-set allows observers to selectively allocate attention to *O*'s when searching for a red *O* among red and green *O*'s and *N*'s (Egeth, Virzi, & Garbart, 1984). Attentional control researchers have debated whether stimulus-based or goal-driven factors play a larger role in guiding attention, but this assumes the balance between stimulus-based and goal-driven factors remains static: an assumption that is simply incorrect

(Vecera et al, 2014). Here I will investigate a determining factor in whether stimulus-based or goal-driven factors direct attention – I will investigate how experience influences where we allocate attention. I will specifically examine this in the context of distraction (for review see Geng, 2014). That is, I will examine how experience influences our ability to prevent distraction.

Prominent models of attentional control posit that a strong, specific bias towards target items is sufficient for avoiding distraction (Bundesen, 1990; Desimone & Duncan, 1995; Wolfe, 1994). This predicts that changing a salient distractor's identity will never influence distraction since attentional control is independent of distractor experience. Here, I will explore the role of distractor experience in attentional control and investigate how we maintain this experience in order to prevent future distraction.

CHAPTER 1. ATTENTIONAL CONTROL

Because more information enters our sensory systems than they can effectively process, our sensory systems have developed mechanisms to prioritize important information. Cognitive researchers named this selective processing of important information attention. Attentional mechanisms must choose what information is important and what information is unimportant, and attentional control researchers examine how attentional mechanisms make this choice. One important question in the attentional control literature is: what information does our sensory system use to make this selection? This question has typically been asked as a dichotomy: does our sensory systems use *stimulus-based factors* (i.e., bottom-up selection) or *goal-driven factors* (i.e., top-down selection) to decide where to allocate attention? I will not attempt to resolve whether stimulus-based or goal-driven factors direct attention. Nonetheless, I will provide a brief description of research supporting stimulus-based and goal-driven accounts of attentional control as this research illustrates how I will measure attentional control.

Stimulus-Based Attentional Control

Advocates of stimulus-based attentional control hypothesize that attention, by default, initially visits the most salient item in a display (for review see Theeuwes, 2010). Possibly the strongest evidence for this comes from the additional singleton paradigm (Theeuwes, 1992). In the additional singleton paradigm, observers search for a shape singleton target among homogenously shaped distractors (e.g., a circle among diamonds) and respond to the orientation of a line embedded within the target (see Figure 1A). On half the trials, all the items are the same color, but on the other half, one distractor is a different color (i.e., a salient color singleton; see Figure 1B). This differently colored item is never the target, so an observer with perfect goal-driven attentional control should never select it. Nonetheless, observers respond slower when the color singleton distractor is present than absent (see Figure 1C). Theeuwes hypothesized that slower RTs when the color singleton is present indicate that observers allocated attention to the color singleton distractor before the target. That is, Theeuwes hypothesized that the color singleton captured attention.

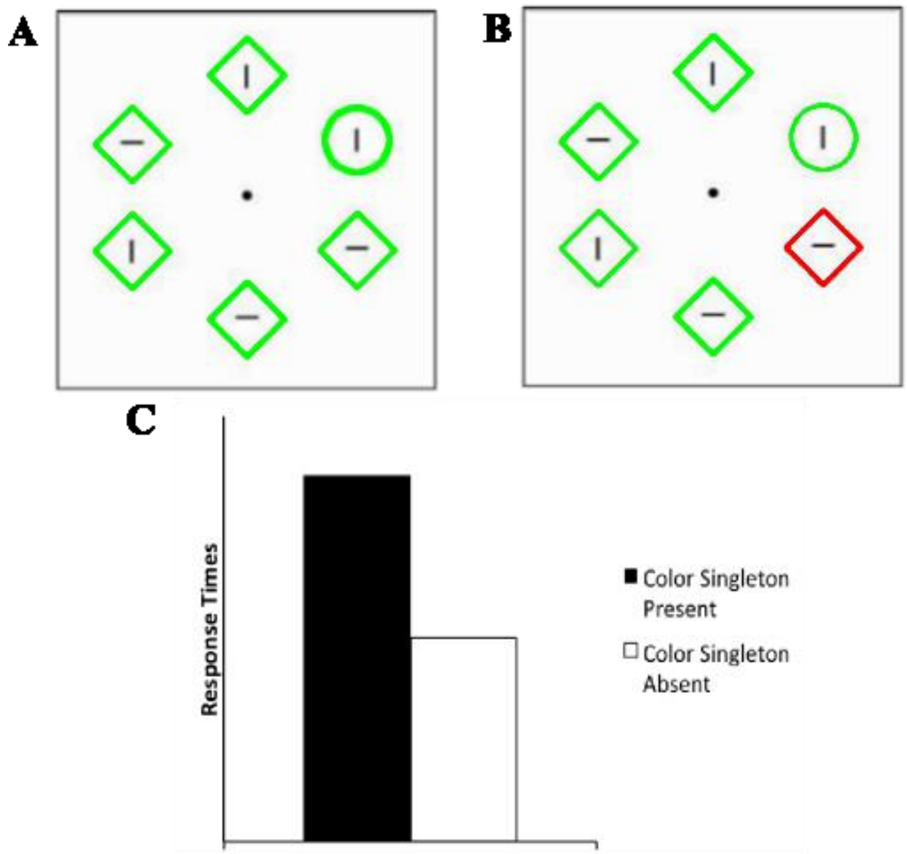

Figure 1. (A) Singleton absent trial from Theeuwes (1992). (B) Singleton present trial from Theeuwes (1992). (C) The typical results found in the additional singleton paradigm.

The attentional capture interpretation of slower RTs when the color singleton is present has not gone unquestioned, however. One alternative account argues that the slower RTs are the result of goal-driven attention, not stimulus-driven attention. In the additional singleton paradigm, observers search for a singleton shape target (e.g., circle among diamonds); in doing so, they might be searching for any difference, or any singleton (i.e., *singleton search mode*, Bacon & Egeth, 1994; Pashler, 1988). If observers were in a 'singleton search mode,' then any singleton—even an irrelevant color singleton—would capture attention because it matched the observers' goals. In favor of a second alternative account, Vatterott, Mordkoff, and Vecera (*in prep*) argued that color singletons do not capture attention in the additional singleton paradigm; instead, color singletons slow RTs by competing for attention with the target (Folk & Remington, 1998). Under this account, although color singletons do not capture attention, color singletons' ability to compete for attention means that RT slowing still measures color singletons' ability to slow attentional allocation to the target. I will refer to color singleton related slowing as distraction. By distraction, I refer to color singletons' ability to slow attentional shifts to the target, and I am agnostic as to whether attentional capture or competition between the target and color singleton generates this slowing.

The stimulus-based account, like many current models of visual attention (Bundesen, 1990; Itti & Koch, 2001; Wolfe, 1994), divides visual processing into two separate stages (Theeuwes, 2010; see Figure 2). In the first stage, the visual system rudimentarily processes all items in the visual field. The second stage has significant capacity limitations so it processes a small portion of the first stage's output. Attentional selection is the decision of what items will enter the second stage. The stimulus-based account claims the input to this second stage depends solely on stimulus-based factors (i.e., stimulus salience). That is, the most salient items from the first stage enter the second stage.

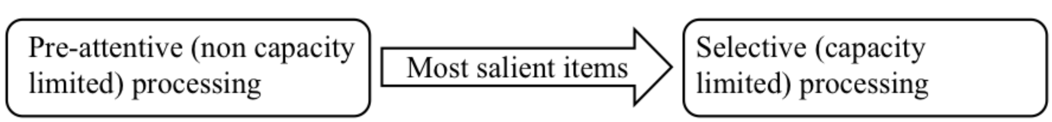

Figure 2. A schematic of the two stage visual processing model posited by the stimulus-based selection account (Theeuwes, 2010).

The importance of salience to this account demands a strong definition of salience. The Merriam-Webster dictionary defines salient as, “very important or noticeable.” Theeuwes defined salient items as, “locations whose local visual attributes significantly differ from the surrounding image attributes” (Theeuwes, 2010, p. 78). Rosenholtz (Rosenholtz, 1999; Rosenholtz, Li, & Nakano, 2007) provided a similar mathematical definition of salience:

$$saliency = (|\mathbf{T} - \mu_D|)' \textstyle\sum_D^{-1}(|\mathbf{T} - \mu_D|)$$

where **T** represents an item’s feature value and μ is the average of all the feature values on feature dimension D across the display. Thus, this equation represents the difference between an item’s feature value and the mean feature value on that dimension across the display, normalized by the amount of variance on this feature dimension. Rosenholtz and Theeuwes’ definition of salience stresses the “noticeable” portion of the definition rather than the “important” portion. This distinction is critical, as salient does *not* mean task-relevant or important in the attentional control literature. I will use the word salient as Theeuwes and Rosenholtz do.

Goal-Driven Attentional Control

Advocates of the goal-driven selection account hypothesize that attention, by default, initially visits the most salient location fitting an observers’ task set. For instance, if an observer’s goal were to find a granny smith apple, they would prioritize green items. When searching for this apple, the goal-driven selection account predicts that observer’s attention will initially visit the most salient green item even if a more salient red item is present. The strongest evidence for the goal-driven selection account comes from the contingent capture task (Folk &

Remington, 1994; Folk et al., 1992). In the contingent capture task, observers either search for a uniquely colored target or an onset target (see Figure 3). Before the onset of the search display, a cue appears at one of the possible target locations. If the cue captures attention, then observers will respond faster when the cue and target appear at the same location than when they appear in different locations (Posner, Snyder, & Davidson, 1980).

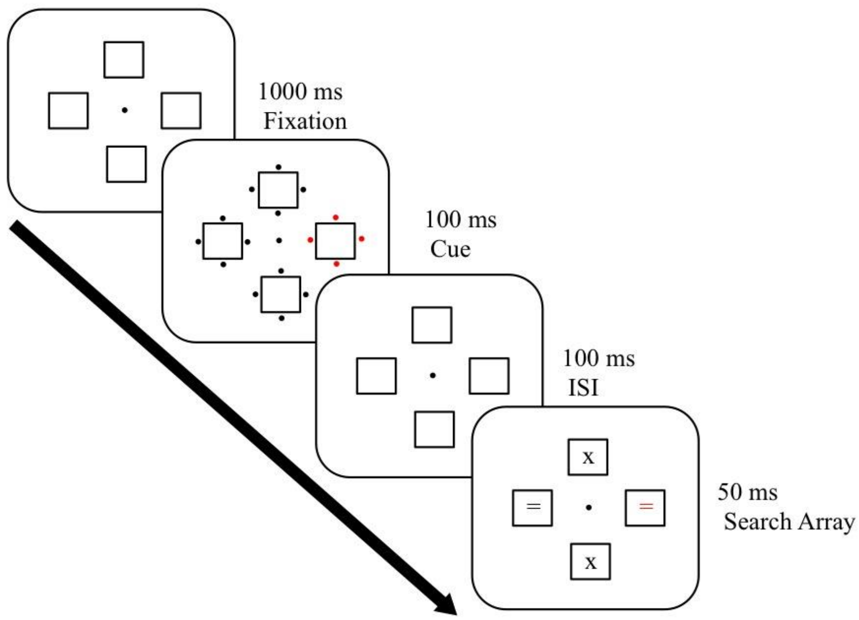

Figure 3. The sequence of events in Folk et al.'s (1992) experiment. This depicts a validly cued trial because the color singleton cue and target appear at the same location.

Folk and colleagues found that observers responded faster when the cue and target appeared at the same location, but only if the cue's feature matched the target. For instance, if the cue was a color singleton, then observers identified a color singleton target faster at the cued location. Thus, only cues fitting observers' task-set captured attention.

The goal-driven selection account, like the stimulus-driven selection account, endorses a two-stage processing model. In this model, the initial stage processes all items in parallel and provides input to a capacity limited stage (see Figure 4). The second stage has significant capacity limitations so it processes a small portion of the first stage's output. Attentional selection is the decision of what items will enter the second stage. The goal-driven selection account claims that the input to the second stage depends on observers' goals *and* the saliency of items in the display. Specifically, the goal-driven account claims that observers' goals gate the second stages' input- only items matching an observer's goals advance to the second stage, and these items are ordered by salience.

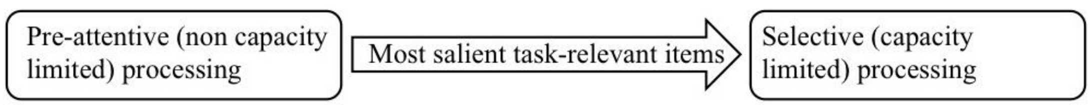

Figure 4. A schematic of the two stage visual processing model posited by the goal-driven selection account.

The goal-driven selection account's reliance on the idea of "goals" demands a strong definition of goals. Theeuwes, an advocate of the stimulus-based account, defined goals as, "an active volitional process" (Theeuwes, 2010, p. 77). This definition highlights the volitional nature of goals, but not all researchers agree that a process must be volitional to be a "goal" (Egeth, Leonard, & Leber, 2010). Instead, advocates of the goal-driven account include non-volitional changes in attentional control due to priming (Maljkovic & Nakayama, 1994), and working memory (Olivers, Meijer, & Theeuwes, 2006) as part of their definition of "goals" (Egeth, Leonard, & Leber, 2010). I will define goals as endogenous task performance changes not explained by information in the current display. This is a strict definition, but I hope the strictness avoids misunderstanding.

CHAPTER 2. MODELS OF ATTENTIONAL CONTROL

In this chapter, I will review a few models of attentional control. There are many models of attentional control, and I will review only a subset of them, but I hope this review clarifies how researchers conceptualize attentional control.

Biased Competition

Biased Competition (Desimone & Duncan, 1995) posits that attentional effects arise when multiple stimuli fall within a single neuron's receptive field. Multiple stimuli within a single receptive field means the neuron might respond to any of these stimuli, rendering that neuron's response ambiguous. Attention resolves this ambiguity by selecting one item in a neuron's receptive field, causing the neuron to selectively respond to this item. Biased Competition posits two mechanisms by which attention chooses what item to select: First, attention preferentially selects items matching those held in working memory. Second, attention preferentially selects salient items. Attention is biased towards items matching those held in working memory because holding an item in working memory causes the cells representing this item to maintain a higher baseline activation level (Luck et al., 1997; Miller & Desimone, 1994). Cells representing items held in working memory also represent these items when they appear in the external world, so their higher baseline activation rate makes these cells more likely to reach threshold and pass their signal along when an item matching the item held in working memory appears. Thus, holding an item in working memory biases attention towards that item.

Neurons respond more vigorously when salient items appear in their receptive field (Allman et al., 1985), biasing attention towards salient items. Interestingly, the visual system also biases processing towards novel items. For instance, Miller, Li, and Desimone (1991) found that cells selective for a certain feature respond more vigorously when that feature is novel than when it is familiar.[1] Thus, even cells selective for a certain feature prioritize that feature more when novel than familiar.

In summary, Biased competition posits that if observers hold a target in working memory, observers will direct attention to that item. Salient, irrelevant items can capture attention, but biased competition offers no explicit mechanisms for biasing attention away from these salient distractors. Observers might be able to bias attention away from items matching those held in working memory (Arita et al., 2012). Biased competition also posits that salient distractors become less salient as they become more familiar, which offers a passive form of learned distractor rejection. Thus, while biased competition does not offer a mechanism for devaluing salient distractors, it does offer a mechanism through which distractors would become less salient across experience.

Theory of Visual Attention (TVA)

Bundesen's Theory of Vision Attention (TVA; 1990; Bundesen et al., 2005; Bundesen et al., 2011) is a mathematical model describing the selection and categorization of visual stimuli. In TVA, visual stimuli race towards selection into short-term memory. TVA describes average data from experimental conditions, so instead of describing the rate at which an item races towards selection, TVA describes the probability that the visual system will select an item

[1] The repetition suppression literature interprets this reduced activity from repeated exposure as more efficient processing (Schacter & Buckner, 1998). This account can be integrated with our salience account if I describe salience as information load (Bruce & Tsotsos, 2009) and repeated displays provide less information than the initial display.

(higher rate = higher probability). The probability that an item reaches short-term memory is described by the following equation:

$$v(x,i) = \eta(x,i)\beta_i w_x / \sum_{z \in S} w_z \tag{1}$$

Here, x refers to a stimulus and i refers to a category. $\eta(x, i)$ refers to the amount of sensory evidence that stimulus x belongs to category i. β_i represents the perceptual bias towards assigning stimulus x to category i. For instance, the visual system might be biased towards categorizing stimulus x as a circle if it has only encountered circles in the recent past. w represents the attentional weighting of a stimulus. Thus, the final part of the equation depicts the weight of stimulus x relative to the weighting of all the other stimuli in display S. In his 1990 paper, Bundesen writes the weighting equation as follows.

$$w_x = \sum_{j \in R} \eta(x,j)\pi_j \tag{2}$$

Here $\eta(x,j)$ represents the probability that stimulus x belongs to category j where j is a possible category in the R set of all possible categories. π_j represents the pertinence of category j. More important items, such as targets, will have a higher pertinence value than distractors. The attentional weight of an item is relative to other items, so if many items appear similar to the target, then many items will receive some pertinence, diluting the bias towards targets and making it less likely for the visual system to select the target. TVA does not allow negative pertinence values, meaning that items cannot be actively inhibited below a particular baseline.

In TVA (Bundesen, 1990), if observers have a sufficiently strong, specific bias towards the target, then this target bias should overwhelm distractors and result in effective target selection. That is, a strong, specific target bias should prevent distraction. This strong, specific bias towards the target would have to be relatively strong compared to the bias towards other items, but TVA's structure does allow the creation of an attentional state that is immune to distraction. Alternatively, the relative attentional weights of items could be constrained such that targets cannot receive such a great attentional weight that they completely dominate the race towards selection.

Bundesen (1990) describes the pertinence value as a free parameter under the experimenter's control, which vaguely suggests that an observer has conscious access to pertinence values like a subject has conscious access to short term memory (STM) contents. Bundesen does not describe whether observers can devalue items. Nordfang, Dyrholm, and Bundesen (2013) used the TVA framework to examine the role of salience in calculating a stimulus' pertinence values. Nordfang and colleagues investigated whether the pertinence values reflected only task settings (such as target or not) or whether the pertinence value also incorporates salience. Nordfang and colleagues found salient items have higher pertinence values, even when completely task irrelevant. Thus, it seems that the visual system is more likely to select a salient item than non-salient items. To account for this, Nordfang and colleagues proposed to change TVA's weight equation to the following:

$$w_x = \mathrm{K}_x \sum_{j \in R} \eta(x, j)\pi_j, \tag{3}$$

the important difference here being the addition of K_x, which represents the feature contrast of stimulus x. Thus, a stimulus with a high feature contrast boosts an items weight irrespective of the item's identity. TVA posits that once we assign appropriate priority to targets, attention will be biased towards these items. TVA, like biased competition, predicts that salient, irrelevant items can capture attention. TVA could allow distractor items to be devalued, but the attentional weights of items must be greater than zero. This means that all items would have to start with an attentional weight greater than zero. Target items would receive greater weights and distractor items would receive lesser weights. TVA has a homunculus set the attentional weights of items, but presumably, these weights could be set by harnessing past experience. Thus, while TVA has the structural ability to devalue distractors, this has never been implemented. Importantly, while TVA can devalue distractors, it postulates that a strong, specific bias towards target items is sufficient for preventing distraction. This makes devaluing distractors unnecessary for preventing distraction.

Logan's Instance Theory (1988; 2002)

The Instance Theory of Attention and Memory (ITAM) advances TVA by positing how the visual system finds the value of $\eta(x, i)$ in equation 1. $\eta(x, i)$ represents the multi-dimensional similarity (multiple features; Shepard, 1957) between a visual stimulus and category x. ITAM posits that this similarity computation proceeds by comparing the visual stimulus to instances in memory (Nosofsky, 1988). The more instances of a category stored in memory, the more precise an observer's representation of that category. The notion that categories are learned over time demonstrates that when searching for a target, not only do observers need to set attentional weights appropriately, but observers also need to learn where to apply these weights. For instance, if an observer is searching for a red circle among green circles, the observer will want to bias selection towards red items, but the observer must also learn exactly what this red stimulus looks like. Otherwise, the observer might mistakenly bias attentional priorities towards a task-irrelevant shade of red. Logan (2002) mentions that observers must learn category representations, but notes that observers' homunculus sets attentional priorities. Thus, ITAM posits that once observers know the task-relevant category and assign a priority to this category, observers will selectively direct attention to task-relevant items. ITAM says nothing about the role of distractor rejection in attentional control. Because ITAM accomplishes attentional control through TVA, ITAM has the structural ability to devalue distractors, but this has never been implemented.

Models of Attentional Control Summary

This brief review of attentional control models demonstrates that researchers posit that distractor rejection plays a rather limited role in attentional control. I do not claim to review all models of attentional control. For instance, I did not review guided search (Wolfe, 1994). I reviewed the models above because they explicitly addressed how attention selects items. No models reviewed here have implemented distractor devaluation. Biased competition posits that items might be devalued as they become more familiar, but this novelty bias is not limited to distractors. Thus, the current implementation of major models of attentional control posit that once observers learn to bias attention towards the target, they will have reached an asymptotic level of attentional control and an irrelevant distractor might distract them, but they will not learn to ignore these items.

CHAPTER 3. ESTABLISHING THE ROLE OF DISTRACTOR EXPERIENCE IN ATTENTIONAL CONTROL

Experiment 1: Learned distractor rejection

Prominent models of attentional control posit that a strong, specific bias towards a target is sufficient for avoiding distraction (Bundesen, 1990; Desimone & Duncan, 1995). For instance, recent implementations of TVA models attentional control as a bias towards targets, which prevents attention from visiting distractors dissimilar from the target. TVA's structure could provide a mechanism for biasing attention away from salient distractors, but this has not been implemented. Thus, TVA predicts that if observers have a strong, specific bias towards a target, then observers will be immune to distraction by dissimilar distractors (but see Nordfang et al., 2013). Bacon and Egeth (1994) confirmed this prediction by demonstrating that salient distractors distract observers if observers have an imprecise target template, but not if observers have a precise target template. In Experiment 1, I investigate whether a strong, specific bias towards target items is sufficient for preventing distraction or if distractor experience also plays a role in attentional control.

In Experiment 1, observers searched for a green target circle among green heterogeneously shaped distractors. Searching for a specific target shape among heterogeneously shaped distractors requires observers to use a specific target template. Thus, if preventing distraction simply requires a strong, specific bias towards targets, these observers should be immune to distraction. Observers began the experiment with a 60 trial training block. After this training block, observers completed four test blocks, each with a differently colored item (i.e., color singleton distractor) present in half the trials. For instance, Block 1 might contain a red distractor item on half the trials and Block 2 might contain a blue distractor item in half the trials. If a strong, specific bias towards a target is sufficient for avoiding distraction, then the training block will provide observers with this strong, specific target bias and salient distractors will never distract observers. If observers also need experience with salient distractors to effectively reject them, then salient distractors will initially distract observers, but observers will quickly learn to ignore them.

Methods

Observers. Sixteen observers (7 females) from the University of Iowa psychology research participant pool participated for partial course credit. All observers reported normal vision and that they were not color blind. The Institutional Review Board approved this study and all observers gave informed consent.

Apparatus. Stimuli appeared on a 17-in. CRT monitor controlled by a Macintosh Mini, using MATLAB and the Psychophysics Toolbox (Brainard, 1997). Observers sat approximately 60 cm from the display.

Stimuli. Six colored shapes appeared equally spaced around the circumference of an imaginary circle centered at fixation with a radius of 4.2°. The fixation point was a small white dot in the center of the screen. Each shape was 2.5° wide and 2.5° tall. Each shape contained a randomly vertical or horizontal white line (0.7 x 1°). Each display had one target circle and five distractors. Each distractor could be a triangle, diamond, or square. Each distractor's identity was randomly assigned on each trial with the only constraint being that the distractors could not all be the same shape. Target position was assigned pseudo-randomly on each trial. During the training block, all items were green. In the test blocks, all the items were green (RGB 0, 255, 0; *singleton absent* trials) in half the trials, and one item was a color singleton (*singleton present* trials) in the other half. The color singleton could be red (RGB 255, 0, 0), yellow (RGB 255, 255, 0), purple (RGB 255, 0, 255), or orange (RGB 255, 150, 0). Color singleton position was assigned randomly from the positions not already occupied by the target. Color singleton color remained constant within blocks, but changed between blocks. A latin-square assigned the color singleton's color in each block.

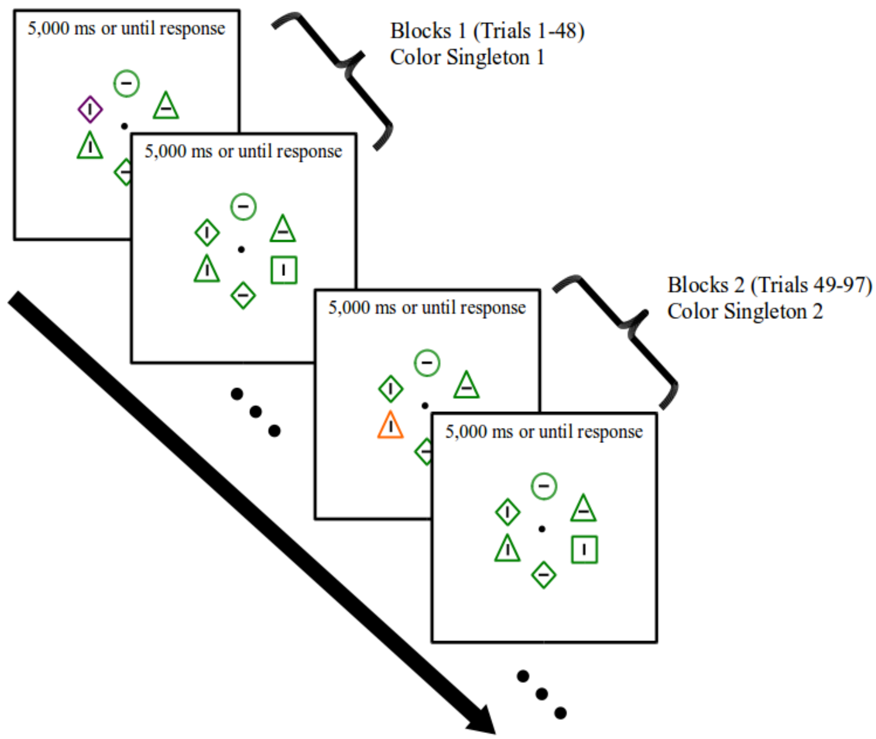

Figure 5. Experiment 1 sequence of events. 1000 ms ITI not depicted. Each block contained a differently colored color singleton distractor.

Design. The general procedure appears in Figure 5. Each trial began with only the fixation point visible. After 1,000 ms, the search display appeared for 5,000 ms or until response. If an observer failed to respond within 5,000 ms, the observer was encouraged to respond faster and the trial was marked as incorrect. A beep informed observers of incorrect responses.

Observers task was to indicate whether the target contained a vertical or horizontal line as quickly and accurately as possible. Observers indicated target line orientation by pressing the “z” or “m” keys.

The experiment began with a 60 trial training block during which participants search for the circle among distractors, and there was never a color singleton distractor. Following the training block, participants completed four 48 trial test blocks. Each block contained a differently colored color singleton. A short, self-paced break preceded each block. Eye-movements were not monitored, but participants were encouraged to maintain fixation.

Results and Discussion

RTs more than five standard deviations above the mean, RTs faster than 300 ms, incorrect RTs, and RTs following an incorrect response were removed from the analyses. The RT trimming eliminated less than 2% of the data.

I have previously depicted learned distractor rejection by binning RTs into the first and second half of each block (Vatterott & Vecera, 2012; see Figure 6). This analysis finds that salient distractors distract observers in the first half of blocks, but not the second half. This implies that the first half of blocks is a pre-learning period, when salient distractors distract observers, and the second half of blocks is a post-learning period, when observers ignore salient distractors. This analysis adequately measures learned distractor rejection, but does not allow one to examine continuous learning. Learned distractor rejection likely occurs continuously rather than in a single stroke of insight, so I will use an alternative analysis that depicts learned distractor rejection as occurring continuously rather than in a single step.

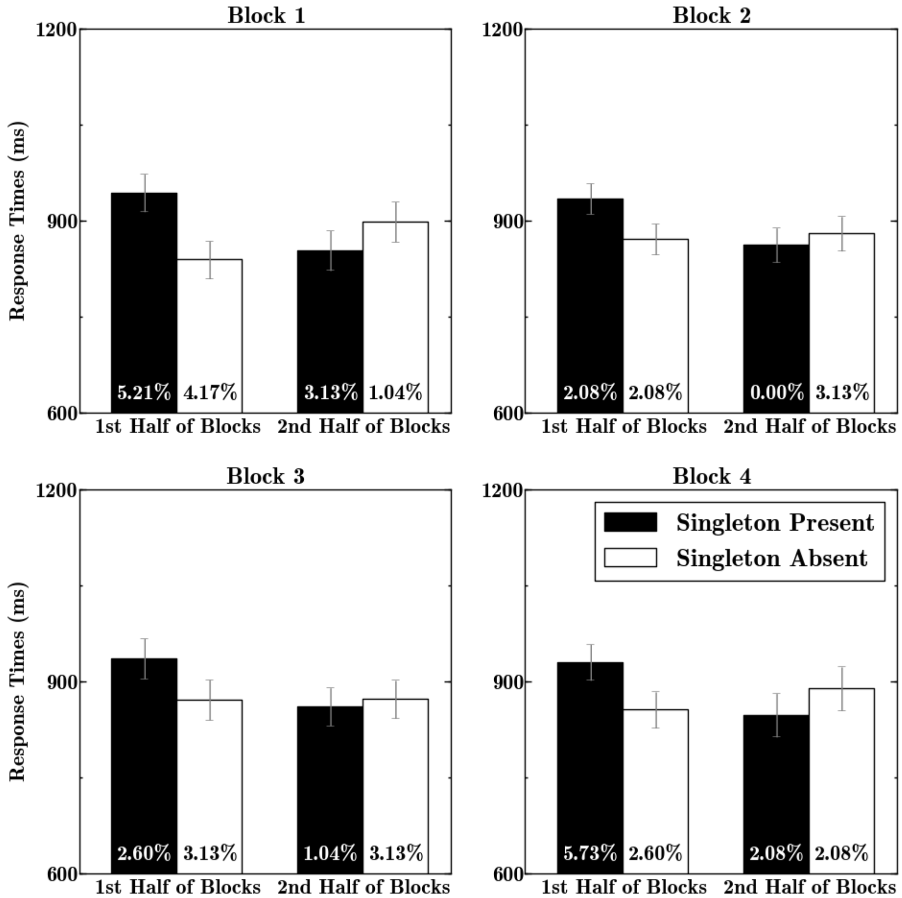

Figure 6. Experiment 1 RTs from each block as a function of singleton presence and block-half. Error rates appear in the base of the bars. Error bars depict 95% confidence intervals (Baguley, 2012; Loftus & Masson, 1994).

To measure learned distractor rejection continuously, I will estimate the parameters of power functions fit to each observer's data from the singleton absent and present conditions (Chun & Jiang, 2003; Vatterott, Mozer, & Vecera, *submitted*). I fit the data to a power function with a free additive constant that allows asymptotic performance to deviate from 0 (Logan, 2002). Specifically. I used the following function:

$$\mathrm{RT} = \beta * \mathrm{Trial\#}^{-\alpha} + \gamma \tag{4}$$

RT is response time and Trial# is the number of experiences an observer has with a particular singleton condition. For instance, an observer's first experience in a singleton present trial during each block is Trial# 1, and an observer's third experience in a singleton absent trial during each block is Trial# 3. Best fits were found using the least squares optimization function from the python package SciPy (van der Walt,Colbert, & Varoquaux, 2011). The parameter β was bound between 0 and 4500. The parameter α was bound between -1 and 1. The parameter γ was bound between 500 and 5000. These parameters were constrained to prevent over fitting of the results. This was particularly important in the singleton absent condition where the relative lack of predictable RT changes encouraged erratic fits. The basic pattern of results remained the same without these boundaries. To avoid local minima, I fed the least squares optimization algorithm 25 randomly generated initial parameter values. I used the parameters defining the best-fitting function after these 25 iterations of the least squares optimization algorithm. Using 100 random initial parameter values did not change the results, suggesting that I avoided local minima.

By fitting the data to a function and using the predicted performance and parameters of this function, performance exhibited by observers in this task is reduced to a format that I can make theoretical predictions about. Specifically, I can measure distraction at the beginning of blocks by comparing predicted performance at the beginning of blocks. I can also measure the amount of learning and rate of learning by examining the β and α parameters, respectively. Finally, I can examine whether asymptotic performance differs between singleton present and absent conditions. If the performance observed in Bacon and Egeth (1994) represents asymptotic performance, γ the parameter representing asymptotic performance, should never differ between singleton present and absent conditions. All of these parameters can be influenced by the entire data set, providing a sophisticated reduction of many RTs into a theoretically tractable format. I do not predict that a power function fits the data better than other functions. I am simply using a power function because of its long history in the learning literature (Newell & Rosenbloom, 1981). Preliminary analyses suggest that the same pattern of results occurs with other functions such as an exponential decay function or power function with an asymptote of zero.

The python packages NumPy (van der Walt,Colbert, & Varoquaux, 2011), IPython (Pérez & Granger, 2007), and Pandas (McKinney, 2010) were also used during data analyses. Data figures were generated using Matplotlib (Hunter, 2007).

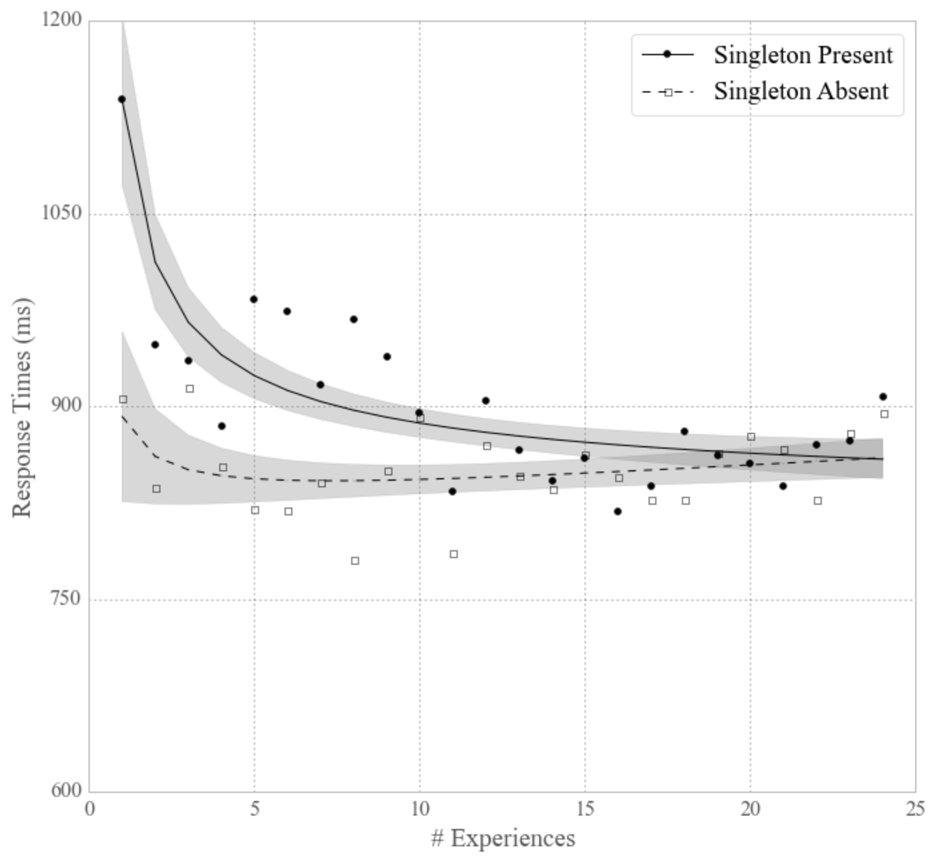

Figure 7. Experiment 1 RTs averaged across blocks as a function of singleton presence and experience. The average RTs predicted by each observers' best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

The parameter β represents the amount of learning between the beginning of the block and asymptotic performance. The parameter α represents how quickly performances changes (learning rate). The parameter γ represents asymptotic performance. I predict that novel color singletons initially distract observers, slowing RTs. The slower initial RTs in singleton-present trials will be captured by greater predicted RTs at the beginning of blocks. I predict that observers will learn to prevent this distraction as exhibited by a larger β parameter in singleton present than singleton absent trials and an α greater than zero in the singleton present condition.

The data and average RTs predicted by each observer's best-fit power function appear in Figure 7. Note that these are average predicted RTs rather than RTs predicted by averaging parameter values together. I represent the data this way because it better represents the average function. Figure 7 uses dots to depict RTs averaged across blocks as a function of singleton presence and experience. For instance, the first time an observer encounters a color singleton in a given block is singleton present trial 1 and this is the first filled dot on the left. Each observer's fit appears separately in Figure 8.

The singleton present functions had an average RMSE of 150. The singleton absent function had an average RMSE of 126. At the beginning of the block, the singleton present functions predicted greater RTs than the singleton absent function, $t(15) = 4.09$, $p < 0.001$, demonstrating that color singletons initially distracted observers. β, the parameter representing the amount of learning between the beginning of the blocks and asymptotic performance was, was greater in the singleton present than absent condition $t(15) = 4.11$, $p < 0.001$. α, the parameter representing learning rate, was significantly greater than 0 in the singleton present condition $t(15) = 4.43$, $p < 0.001$, but not in the singleton absent condition, $t < 1$. γ, the parameter representing initial performance did not differ between the singleton absent and present condition, $t < 1$. These analyses demonstrate that RTs were initially much higher in the singleton present condition than the singleton absent condition and RTs decreased more in the singleton present condition than the singleton absent condition. The color singleton initially distracted observers, but observers quickly learned to ignore the salient distract. Observers expressed this learning in the singleton present trials, but not the singleton absent trials.

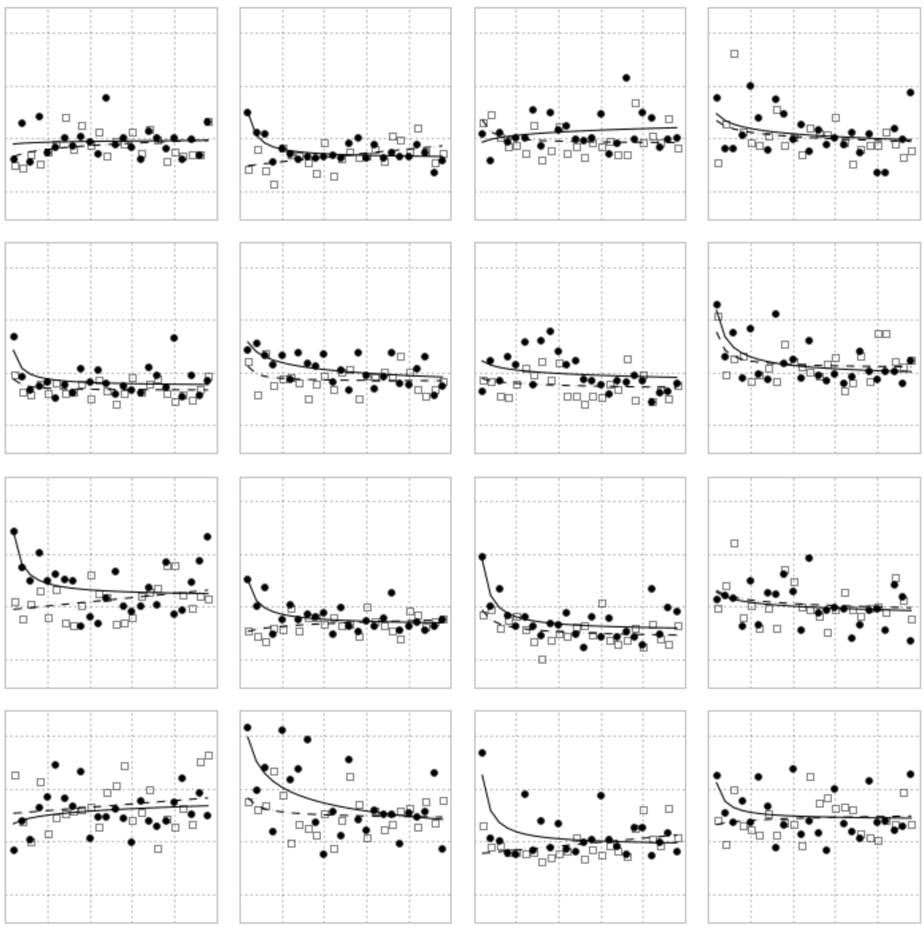

Figure 8. Each Experiment 1 observers' RTs as a function of singleton presence and experience averaged across blocks. The filled in dots represent singleton present trials. The open square represent singleton absent trials. The solid line depicts the best-fit power function for singleton present trials. The dotted line depicts the best-fit power function of singleton absent trials.

Novel distractors distracted observers even though observers had a strong, specific bias towards the target. These results demonstrate that a strong, specific bias towards the target is insufficient for preventing distraction. Instead, preventing distraction also requires that observers learn to ignore the salient distractors. These results suggest that models of attentional control should be modified such that there is an inherent bias towards salient, novel items, which decreases as observers gain experience with the items, possibly by actively inhibiting features associated with the distractor (Moher et al., 2014).

Experiment 1 perfectly confounds block change and color change. That is, every time observers experienced a new color singleton, observers also recently took a break. This break might have loosened observers' attentional control settings and rendered observers more vulnerable to distraction. This possibility is unlikely as observers maintain attentional control settings for as long as a week (Leber et al., 2009). Nonetheless, I demonstrated that novel color singletons distract observers when novel color singletons are not preceded by a break (Vatterott & Vecera, 2012), indicating that color singleton novelty produced the distraction observed here. I also recently demonstrated that distraction does not increase following a break, indicating that breaks do not change distraction (Vatterott et al., *under revision*).

The power function fitting analysis used in Experiment 1 demonstrates that learned distractor rejection, like skill learning (Newell & Rosenbloom, 1981), occurs continuously. The power law of practice describes learning in tasks from cigar rolling (Crossman, 1959) to categorization (Nosofsky & Palmeri, 1997). My finding that the power law of practice appears consistent with learned distractor rejection suggests that the same learning principles that govern tasks such as cigar rolling also govern the unrelated task of distractor rejection. While I used a power function to describe the data here, I do not claim the power function fits the data better than other functions (although it might). I simply used to the power function because of its long history of reliably describing learning. I could have fit the data with other functions, such as an exponential decay function. Just as multiple functions can describe skill learning (Heathcote, Borwn, & Mewhort, 2000); multiple functions could very well describe learned distractor rejection. Importantly, learned distractor rejection, like skill learning, occurs continuously across experience. This indicates that we do not simply instate an attentional control state by selecting a search mode (Bacon & Egeth, 1994) or by choosing a target template (Desimone & Duncan, 1995). Instead, attentional control is gradually honed across experience (Logan, 2002; Chun & Jiang, 2003; Cosman & Vecera, 2014). The attentional control system should not be thought of as a monolithic cognitive system ruled by either stimulus-driven or goal-driven inputs. Instead, the attentional control system constantly adapts to the task, changing the priorities of stimulus and goal-driven inputs (Vecera, Cosman, Vatterott, & Roper, 2014). Thus, the attentional control system adaptively configures itself to the current task environment just as motor control systems gradually configure to the task of cigar rolling.

TVA (Bundesen, 1990) posits that when observers bias attention towards the target, distraction becomes less likely. In the case of an extremely large target bias, observers should be immune to distraction. Experiment 1 demonstrates that when observers search for a specifically, shaped item, the target bias is not sufficiently strong to render observers immune to distraction. Instead, novel, salient distractors initially distract observers before observers learn to ignore them. TVA could implement this learned distractor rejection by limiting the relative bias towards target items such that observers are not immune to distraction. TVA could also implement a mechanism by which observers devalue the attentional weights of salient distractors, preventing attention from visiting these items.

To guide attention, goals must be instantiated in the brain. Biased competition (Desimone & Duncan, 1995) proposed that we maintain goal representations in working memory. Cosman and Vecera (2014) argued that while some goal representations may be held in working memory, other goal representations might be gradually, implicitly learned. Cosman and Vecera's account stresses the importance of experience in instantiating goal representations (also see Vecera et al., 2014). In Experiment 1, I demonstrated that distractor experience is necessary for effective distractor rejection, but the exact memory system used to represent distractor experience remains uncertain. In the following chapter, I will discuss different demonstrations of experience's influence on attentional control and the memory systems underlying these effects.

CHAPTER 4. EXPERIENCE AND ATTENTIONAL CONTROL

Both stimulus-based and goal-driven factors contribute to attentional control. Attentional control researchers have begun to recognize experience's role in balancing stimulus-based and goal-driven factors contribution to attentional control (Awh et al., 2012). For instance, observers identify salient, color singleton targets faster if the color of this target repeats than if it changes (i.e., priming of pop-out; Maljkovic & Nakayama, 1994). This demonstrates that even when observers have a strong stimulus-based bias towards the target, experience strengthens this bias. Thus, priming of pop-out demonstrates experience's ability to changes the efficacy of attentional control (Lee, Mozer, & Vecera, 2009). The exact mechanism underlying experience's influence on attentional control remains uncertain. This chapter reviews different manifestations of experience's influence on attentional control.

Visual Working Memory Contents and Attentional Control

Visual working memory (VWM) refers to information actively held by our cognitive systems. VWM representations can be differentiated from visual long-term memories (VLTM) based on three properties of VWM (Luck, 2008). First, VWM representations are actively maintained, while VLTM representations passively persist over extensive periods. This requirement of active maintenance limits the length of time over which VWM representations can persist. In theory, VWM representations can persist indefinitely, but this would consume resources needed for other tasks. Second, VWM representations are rapidly encoded (Vogel, Woodman, and Luck, 2006), while VLTM representations accumulate over much longer periods. Third, VWM can only simultaneously maintain 3-4 items (Phillips, 1974), while there is no limit to the number of simultaneously maintained VLTM representations (Brady, Konkle, Alvarez, & Oliva, 2008).

Attention controls what items enter VWM (Sperling, 1960), but Biased Competition (Desimone & Duncan, 1995) posits that VWM representations also influences where we allocate attention (also see Duncan & Humphreys, 1989). According to Biased Competition, when observers search for a target, observers hold a *target template* in VWM, where this target template is an example of observers' target. For example, when searching for a granny smith apple, observers will place an imagined granny smith apple in VWM. This target in VWM excites neural populations representing the visual features defining the target (Postle, 2006). In the granny smith apple example, the neural populations representing green and round will be more active. Because these populations have a greater baseline activation level, they are more likely to reach threshold and pass on their signal (Chelazzi et al., 1993). Therefore, Biased Competition predicts that attention is more likely to select items matching those held in VWM than items not held in VWM.

Downing (2000) tested Biased Competition's prediction that attention is more likely to select items held in VWM. In Downing's experiment, observers first saw a face, which they held in VWM for a later recognition test. Following the face, two faces briefly appeared to the left and right of fixation. One of these two faces matched the face held in VWM. Immediately after the briefly appearing faces, a search target appeared in one of the positions recently occupied by the faces. Downing found that observers responded faster when the target appeared at the position previously occupied by the face matching the face held in VWM. Thus, just like an abrupt onset draws attention to a location in a Posner cueing task (Posner et al., 1980), an item matching an item held in VWM also draws attention to its location.

Soto et al (2005) also investigated the role of VWM items in attentional control. In their task, observers first saw a colored shape, which they held in VWM for a later recognition test. After this shape disappeared, observers searched for a tilted line among vertical lines, and reported the target line's tilt (left or right). Importantly, colored shapes framed each line, and the outlined shapes could match the to-be-remembered item. Even though the to-be-remembered item infrequently matched the shape framing the target (making it task irrelevant), observers reported the target line's tilt faster when the to-be-remembered item framed the target than when it framed a distractor. These results demonstrate that attention preferentially selects items matching those in VWM, even if these items are task irrelevant (also see Olivers et al., 2006).

Woodman and Luck (2007) investigated the automaticity by which VWM representations guide attention. A strong version of Biased Competition posits that all VWM representations bias attention towards matching items, but it is possible that only task-relevant VWM representations bias attention. To test the automaticity of VWM representations' ability to bias attention towards matching items, Woodman and Luck instructed observers to hold a colored square in VWM while searching for a colored landolt-c target. On half the trials, one of the distractors was the same color as the VWM item. The target's color never matched the VWM item, providing a strong incentive to bias attention away from items matching those in VWM. Observers responded faster when a distractor's color matched the VWM item than when no items matched the VWM item (also see Han & Kim, 2009). This demonstrates that observers can use VWM representations to actively bias attention away from irrelevant objects, suggesting that observers might bias attention away from irrelevant items by holding *templates for rejection* in VWM (Arita, Carlisle, & Woodman, 2012). Thus, VWM representations do not automatically bias attention towards matching items.

Arita, Carlisle, and Woodman (2012) recently demonstrated that cueing observers to ignore a particular color biases attention away from items with this color. In this experiment, observers viewed a single colored square, indicating which color the target would *not* be. Following this colored square, a search array appeared; half the items were the negatively cued color and the other half were not related to the previous cue. Cueing the color the distractor color improved performance, demonstrating that observers can maintain a *template for rejection* in VWM. This template for rejection biases attention away from irrelevant distractors.

If VWM maintains target templates, then filling VWM with non-target items should prevent observers from using a target template to guide search, and searching without a search template will reduce search efficiency. Woodman, Vogel, and Luck (2001) tested this prediction by instructing observers to hold four items in VWM while completing a challenging search task. Holding items in VWM did not reduce search efficiency (no change in search slope), demonstrating that observers did not need use VWM to maintain a target template. Woodman, Luck, and Schall (2007) noted that search with a full VWM might have remained efficient because the target item in Woodman, Vogel, and Luck (2001) never changed. Because the target item never changed, the target template might have been maintained in VLTM instead of VWM. To test this idea, Woodman, Luck, and Schall measured search efficiency when observers searched for a repeated and changing target while maintaining a full VWM load. Observers with a full VWM load searched less efficiently when the target changed each trial than when the target remained consistent across the entire experiment. These results demonstrate that VWM only stores a target template when the target frequently changes.

Carlisle et al (2011) measured the amplitude of the CDA ERP component across different amounts of experience searching for a target. The CDA is an ERP component indexing the number of items held in VWM (Vogel & Machizawa, 2004). For example, the CDA has a greater amplitude when observers hold four items in VWM than when observers hold two. Carlisle et al found when observers began searching for a new target, the CDA had a greater amplitude than after observers had searched for this same target over a couple consecutive trials. In fact, the CDA completely disappeared after seven consecutive trials of searching for the same target. These results demonstrate that when we begin searching for a new target, we load a representation of this target into VWM, but we offload this target template to other memory stores as we gain experience searching for the target.

Priming of Pop-out

Priming of Pop-out (PoP) refers to the finding that observers identify a pop-out search target faster if the target-defining feature repeats from one trial to the next (Maljkovic & Nakayama, 1994). For example, Maljkovic and Nakayama instructed observers to search for an oddly colored diamond and to respond to whether the diamond had a notch missing on the left or right side. Observers responded faster when the color of the oddly shaped diamond repeated than when it changed. This response speeding is interesting because searching for oddly colored items can be accomplished solely based on stimulus-based factors. Nonetheless, the repetition advantage of PoP suggests that repeating a target item’s features strengthens the attentional bias towards this item (Becker, 2008).

Tulving and Schacter (1990) posited that perceptual priming occurs via a general property of *perceptual representation systems*. Perceptual representation systems consist of the cognitive systems used to process perceptual features, like color. Each system has a short-term memory of what it recently processed, and this memory expedites processing of repeated information. Therefore, when the visual system repeatedly processes the same items, these items are processed faster, facilitating attentional allocation to the target.

Kristjánsson et al (2007) found a decrease in the visual cortex BOLD response when target features repeated compared to when they changed, suggesting that PoP facilitates processing of repeated features. Additionally, when the color of the target repeated, activity in V4, a brain area involved in color processing, decreased. This suggests that repeating the target's color enabled V4 to process the target's color more efficiently. Presumably, this decreased brain activity, led to the faster responses when the target repeated.

Lee, Mozer, and Vecera (2009) investigated whether PoP depends on VWM resources. VWM and PoP both rely on changes in visual cortex (Postle, 2006; Kristjánsson et al., 2007), so the two might share a common memory system. Observers completed a PoP task with a full VWM load. The VWM load did not change PoP. In a subsequent experiment, Lee et al found interleaving an irrelevant task between PoP trials did interfere with PoP. Thus, while PoP is mutable, PoP does not rely on the same memory system as VWM.

Contextual Cueing

If an observer searches for a T among Ls, performance improves as the experiment progresses. Contextual cueing refers to the finding that performance improves even more if the configuration of targets and distractors repeats across the experiment (Chun & Jiang, 1998). Chun and Jiang originally demonstrated contextual cueing by having observers search for a T among Ls; unannounced to the observers, the configuration of Ts and Ls repeated in half the trials. Observers found the target faster when the configuration of Ts and Ls repeated. These results demonstrate that observers encode the configuration of targets and distractors and use these memories to expedite attentional shifts to the target.

Amnesiac patients with hippocampal damage exhibit improved performance as they gain experience with a search task, yet these patients do not demonstrate contextual cueing (Chun & Phelps, 1999). The hippocampus' role in long-term memory (Cohen & Eichenbaum, 1993) and hippocampal patients' failure to exhibit contextual cueing strongly implies that contextual cueing relies on long-term memory. Contextual cueing, like other forms of long-term memory (Brady & Chun, 2007; Olson & Chun, 2002), is specific to the context in which observers learn it (Brooks, Rasmussen, & Hollingworth, 2010). For instance, Brooks et al found that contextual cueing occurred when a consistent context surrounded the repeated display configuration. Contextual cueing did not occur if a novel context surrounded the repeated display configuration. These results demonstrate that contextual cueing relies on long-term memory systems.

Vickery et al (2010) found that holding task irrelevant items (filling visual working memory) or spatial locations (filling spatial working memory) did not change the size of the basic contextual cueing effect. This demonstrates that contextual cueing does not rely on either VWM or spatial working memory resources. This finding has been questioned though, as Manginelli et al (2013) and Travis, Mattingley, and Dux (2013) found spatial working memory reduced contextual cueing (but see Annac et al., 2013). This reduced contextual cueing might have occurred because the spatial working memory load prevented efficient search (Woodman & Luck, 2004), which prevented observers from exhibiting contextual cueing (Annac et al., 2013). Together, the contextual cueing literature supports the idea that contextual cueing relies on long-term memory rather than working memory systems.

Spatial Probability Cues

Observers identify targets appearing in probable locations faster than targets appearing in improbable locations (Shaw & Shaw, 1977). For example, observers to identified letters that could appear at eight different locations. Targets appeared at some locations more frequently than others. Observers identified targets faster at the frequent locations than the infrequent locations, indicating that observers learned the association between targets and particular locations. This advantage at probable locations is called spatial probability cueing.

In Shaw and Shaw's study (1977), a target was more likely to appear in the same location on consecutive trials if the location was probable than improbable. Thus, like in PoP (Maljkovic & Nakayama, 1994), repetition priming might have driven the faster responses when the target appeared at a probable location. Walthew and Gilchrist (2006) found observers did not exhibit spatial probability cueing when target location did not repeat within four trials. These results suggest that spatial probability cueing occurs by priming perceptual representation systems.

Geng and Behrmann (2005) posited that if spatial probability cueing effects arise from repetition priming, then performance should improve by equivalent amounts when targets repeat at probable and improbable locations. Conversely, if spatial probability cueing effects arise from more than simple repetition priming, then target repetitions at probable locations should improve performance more than target locations at improbable locations. Observers exhibited a greater repetition benefit at probable locations than improbable locations. Jiang et al (2013) found that spatial probability effects persisted even when the training and testing sessions were separated by at least a week. Together, these studies suggest that while target repetitions may be critical for learning the statistical structure underlying the task, long-term memory systems, rather than perceptual representation systems, maintain the representations that lead to probability cueing.

Experience and Search Strategy

Attentional control is the choice of where to allocate attention, and observers use different strategies to make this choice. For instance, in some situations, observers use a *singleton detection mode* (Pashler, 1988) strategy in which they search for any oddball items. In other situations, observers use a *feature search mode* (Bacon & Egeth, 1994) strategy in which they search for items with specific features. Bacon and Egeth found that color singleton distractors do not distract observers when displays encourage observers to use a feature search mode strategy. Color singleton distractors do distract observers when displays encourage observers to use a singleton detection mode strategy. Leber and Egeth (2006) trained one group of observers to use a feature search mode strategy and another group to use a singleton search mode strategy. After the training, both groups searched through displays that would have ordinarily encouraged a singleton search mode strategy. Leber and Egeth found that the color singleton distractor distracted observers if the observers had been previously trained to use a singleton detection mode strategy, but not if the observers had been trained to use a feature search mode strategy. Thus, experience using a search strategy changes our current search strategy, modulating our ability to ignore distractors.

Leber, Kawahara, and Gabari (2009) found that training with either a singleton detection mode strategy or a feature search mode strategy perseverated into a testing session over a week after the training session. This extended perseveration demonstrates that long-term memory stores represent experience with a given search strategy. Cosman & Vecera (2013) trained MTL patients to use feature search mode and found that feature search mode did not perseverate in the MTL patients. Because the MTL is necessary for the formation of new long-term memories, this result strongly demonstrates that search mode perseveration relies on long-term memory systems.

Zehetleitner, Goschy, and Müller (2012) trained observers to use singleton detection mode or feature search mode, as Leber and Egeth (2006) had previously done. Unlike Leber and Egeth, Zehetleitner et al changed the salient distractor's color between the training and test phases. Changing the color resulted in salient distractors distracting observers trained to use singleton detection mode and observers trained to use feature search mode. These results suggest that dissimilar training and test phases reduce the likelihood that observer will perseverate their search strategy. A mismatch between the training and test phase pushes observers towards using the default strategy of singleton detection mode (also see Kawahara, 2010). Cosman and Vecera's (2012) associated displays that encourage feature search mode with one context (e.g., forests) and displays encouraging singleton detection mode with another context (e.g., cities). Observers subsequently used the search strategy associated with a particular context after the association between context and search strategy was removed. Thus, perseveration of search strategy relies on a context match between training and test phases, similar to long-term memory's reliance on a contextual match between training and test phase (Cohen & Eichenbaum, 1993). This similarity suggests that search strategy perseveration relies on long-term memory systems.

Reward and Attentional Control

The goal of attentional control is to allocate attention in the most environmentally efficient fashion, maximizing reward. Raymond and O'Brien (2009) found that observers recognized previously rewarded items better than unrewarded items. This demonstrates that observers' preferentially encoded rewarded items. Anderson, Laurent, and Yantis (2011) instructed observers to find a target circle, which was one of two possible target colors. Each correct response led to a monetary reward, and one target color was associated with higher rewards than the other color. After this training phase, observers transferred to the test phase in which they searched for a green circle among diamonds. On half the trials, one of the diamonds was a color singleton and its color was previously associated with high or low rewards. Color singletons previously associated with high rewards distracted observers more than distractors previously associated with low rewards. Thus, experience biases attention towards previously rewarding items.

Anderson and Yantis (2013) found that color singletons previously associated with a high reward distract observers more than color singletons previously associated with a low reward even when the reward association had not been reinforced in over half a year. The persistence of this reward bias strongly suggests that these reward associations rely on long-term memory systems.

Controlled and Automatic Information Processing

Task experience often reduces task difficulty. For example, driving on a busy street is more difficult for a novice than an experienced driver. This reduced difficulty might arise from a change in the attentional strategy used by novices and experts. Schneider and Shiffrin (1977; Shiffrin & Schneider, 1977) posited two distinct forms of visual search. In the first form, controlled search, observers hold a target template in short term memory, and serially compare items to this target template. Because the comparisons occur serially, observers exhibit poorer performance at larger set sizes (Experiment 1; Schneider & Shiffrin, 1977). In the second form, automatic detection, observers passively hold a target template in long-term memory and compare items to this target template in parallel. Because the comparisons occur in parallel, increasing the display size has little effect on performance. Automatic detection only arises after extended practice (about 1000 trials) with the same set of targets (Experiment 1; Shiffrin & Schneider, 1977). These results demonstrate that extended practice searching for the same target enables observers to select the target with extreme efficiency. Shiffrin and Schneider (1977) found that training observers to search for the same target among the same distractors for over 1000 trials and then switching the target and distractors (former distractors become targets) dramatically reduces performance. In fact, performance is worse than when observers first begin the task. This suggests that training with the same target and same distractors causes observers to prioritize the target, deprioritize the distractors, or both. When targets and distractors subsequently change, these experience driven priorities then negatively influence task performance. The extensive practice needed to develop automatic detection strongly suggests that experience creates attentional templates in long-term memory.

Experience and Attentional Control Summary

Three different memory systems seem to influence attentional control. First, VWM influences attentional control (Downing, 2000). For example, the visual system can bias attention towards (Downing, 2000) or away from (Woodman & Luck, 2007; Arita et al., 2012) items matching VWM contents. Second, perceptual representation systems influence attentional control. For instance, observers allocate attention to targets faster if target features repeat from one trial to the next (Maljkovic & Nakayama, 1994). Third, visual LTM influences attentional control. For example, observers retain biases towards previously rewarded items for over six months (Anderson & Yantis, 2013).

Identifying what particular memory systems give rise to different experience driven attentional effect can be difficult, but this review identified a few critical experiments for solving this problem. If an experience driven effect relies on VWM, then filling VWM with extraneous items should obliterate the experience-driven effect. For example, Lee, Mozer, and Vecera (2009) found filling observers' VWM did not change PoP, which suggests that VWM is not involved in PoP. If an experience driven effect relies on perceptual representation systems, then the local task statistics should entirely explain the effect. For example, Geng and Behrmann found that performance improves more when targets repeat at likely locations than unlikely locations, demonstrating that spatial probability cueing cannot be explained with only local statistics. If an experience driven effect relies on LTM, then the effect should persist over extended delays. For example, Jiang and colleagues (2013) found observers still exhibited learned spatial probability effects after an extended period of disuse, suggesting that LTM representations support spatial probability cueing. I will use these methods to investigate what memory system underlies learned distractor rejection.

CHAPTER 5. DETERMINING THE MEMORY SYSTEM UNDERLYING LEARNED DISTRACTOR REJECTION

Experiment 2: Preserved learned distractor rejection after an extended period of disuse

Many examples of experience's influence on attentional control often rely on LTM. For instance, contextual cueing (Chun & Phelps, 1999), spatial probability cueing (Jiang et al., 2013), experience driven search strategies (Cosman & Vecera, 2013), and reward's influence on attentional control (Anderson & Yantis, 2013) all rely on LTM. In Experiment 2, I test whether learned distractor rejection also relies on LTM. In Experiment 2, observers will perform the same task as Experiment 1, except they will complete two blocks on one day and then complete another two blocks no less than 48 hours later. One of the two salient distractors from the second session will be the same color as a salient distractor in the first session. This repeated color singleton will be familiar when it appears in the second session. If learned distractor rejection relies on LTM, then observers will immediately reject the familiar salient distractor during the second session. If learned distractor rejection does not rely on LTM, then observers will need to relearn to ignore the familiar color singleton. Regardless of whether learned distractor rejection relies on LTM or not, observers will have to learn to ignore the novel color singleton in the second session.

Methods

Observers. Twenty-four observers (15 females) from the University of Iowa psychology research participant pool participated for partial course credit. All observers reported normal vision and that they were not color blind. The Institutional Review Board approved this study and all observers gave informed consent. None of the observers in Experiment 2 had participated in other experiments presented here. More observers participated in Experiment 2 because the critical conditions in Experiment 2 are only one block each.

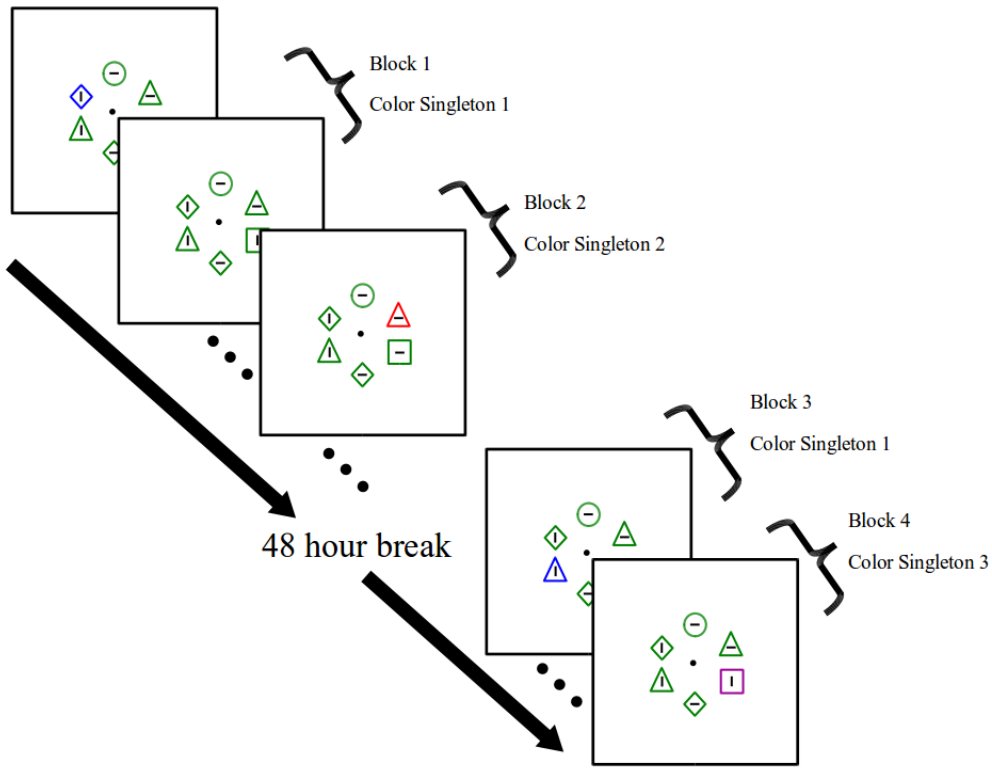

Figure 9. Sequence of events for Experiment 2. 1000 ms ITI not depicted.

Stimuli and Design. This experiment follows Experiment 1's methods except for the following changes (see Figure 9). First, the experiment has four test blocks, but observers complete two test blocks on Day 1 and another two test blocks on Day 2. Observers also completed twelve practice trials without a color singleton on Day 2. In one of the Day 2 test blocks, a color singleton color repeated from Day 1. The other test block contained a novel color singleton color. Color singleton order was counter balanced across observers. The familiar color singleton appeared in the first Day 2 block for half the observers. It appeared in the second Day 2 block for the other half of observers.

Results and Discussion

I used the same exclusion criteria as Experiment 1. The RT trimming excluded less than 0.2% of the data.

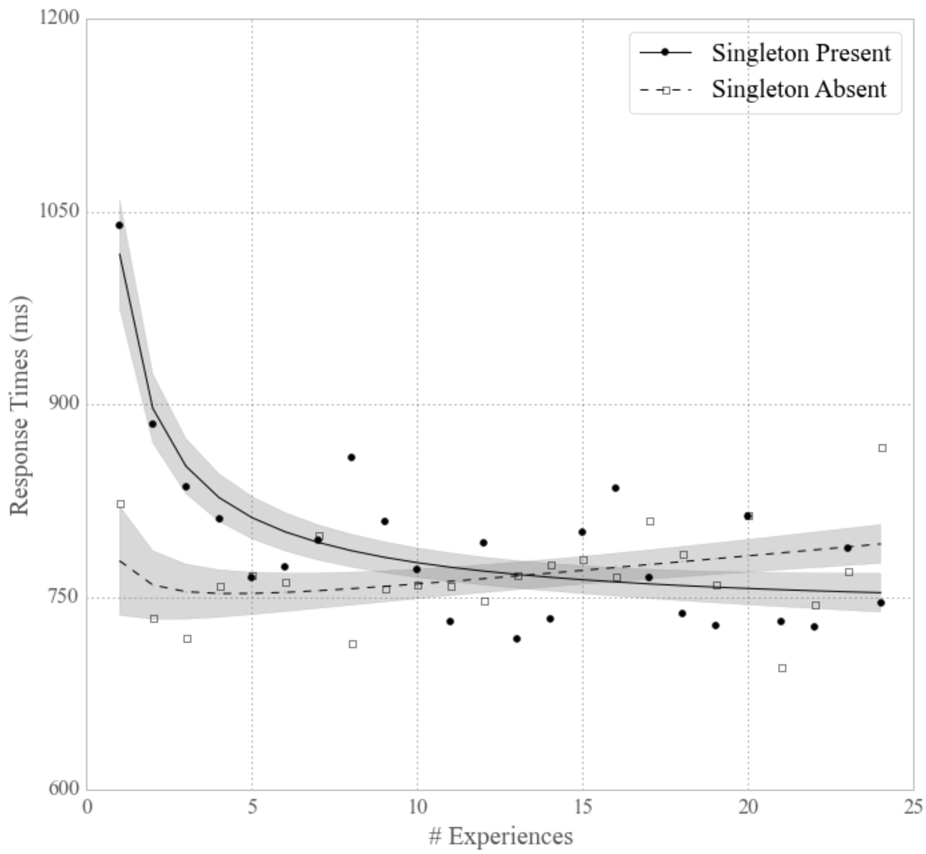

Figure 10. Experiment 2 RTs from the Day 1 test session as a function of singleton presence and experience. The average RTs predicted by each observer's best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

The data and average predicted RTs of the best-fit power functions from the Day 1 test blocks appear in Figure 10. The singleton present functions had an average RMSE of 149. The singleton absent function had an average RMSE of 148. At the beginning of the block, the singleton present functions predicted greater RTs than the singleton absent function, $t(23) = 6.02$,

$p < 0.001$, demonstrating that color singletons initially distracted observers. β, the parameter representing the amount of learning between the beginning of the block and asymptote, was greater in the singleton present condition, $t(23) = 4.86$, $p < 0.001$. α was greater than zero in the singleton present condition, $t(23) = 3.28$, $p < 0.005$, but not in the singleton absent condition, $t(23) = 1.15$, $p > 0.26$, demonstrating that observers learned to ignore the color singleton. γ, the parameter representing asymptotic performance did not differ between the two condition, $t < 1$. Thus, observers replicated Experiment 1, in the Day 1 test blocks.

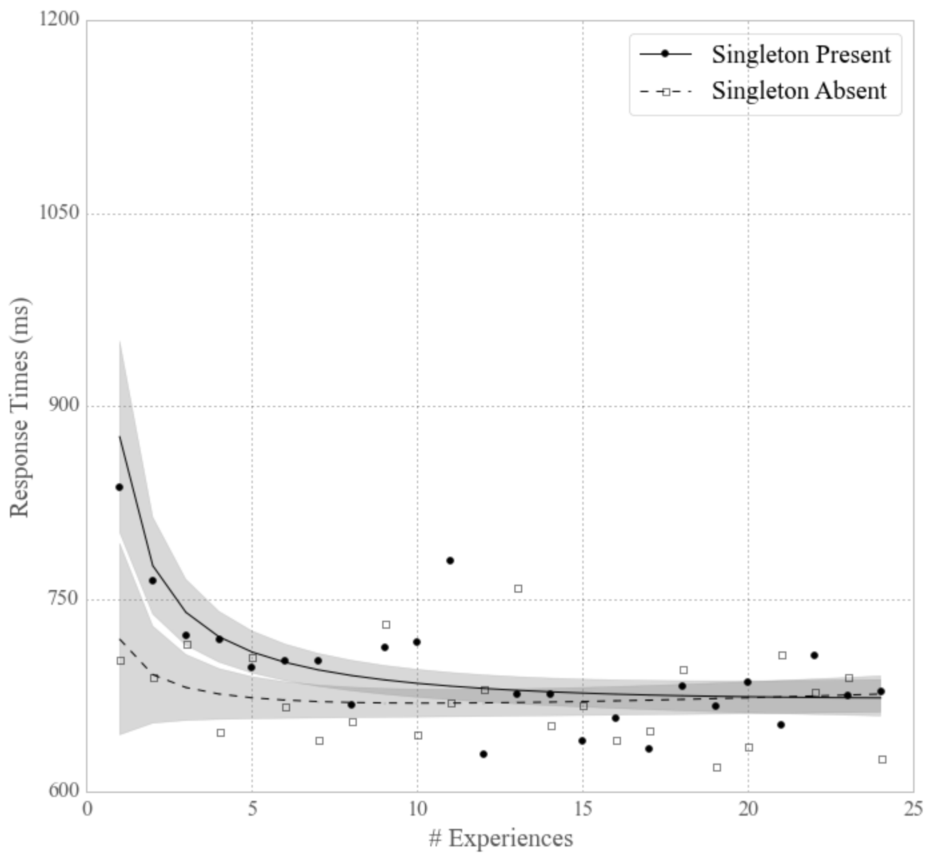

Figure 11. Experiment 2 RTs from the novel singleton condition as a function of singleton presence and experience. The average RTs predicted by each observer's best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

Next, I analyzed data from when a novel color singleton appeared in the second day. The data and average predicted RTs of the best-fit power functions appear in Figure 11. The singleton present functions had an average RMSE of 142. The singleton absent function had an average RMSE of 134. At the beginning of the block, the singleton present functions predicted greater RTs than the singleton absent function, $t(23) = 2.26$, $p < 0.05$, demonstrating that color singletons initially distracted observers. β, the parameter representing the amount of learning between the beginning of the block and asymptote, was greater in the singleton present condition, $t(23) = 2.19$, $p < 0.05$. α was not greater than zero in either the singleton present condition, $t(23) = 1.07$, $p > 0.29$, nor the singleton absent condition, $t < 1$. Thus, the Day 2, novel color singleton test block roughly replicates Experiment 1. The alpha parameter was not significantly different from 0 in the singleton present condition, but observers overall exhibited a pattern of distraction early in the block and little distraction late in the block.

Finally, I analyzed data from when a familiar color singleton appeared in the second day. The data and average predicted RTs from the best-fit power functions appear in Figure 12. The singleton present functions had an average RMSE of 133. The singleton absent function had an average RMSE of 151. At the beginning of the block, observers responded equally fast when the singleton was present and absent, $t(23) = 1.85$, $p > 0.07$. β, the parameter representing the amount of learning between the beginning of the block and asymptote, did not differ between the singleton present and absent conditions, $t(23) = 1.47$, $p > 0.15$. α was not greater than zero in either the singleton present condition, $t < 1$, nor the singleton absent condition, $t(23) = 1.59$, $p > 0.12$, demonstrating that observers consistently ignored the familiar, salient distractor across the block. γ did not differ between the singleton present and absent condition, $t < 1$.

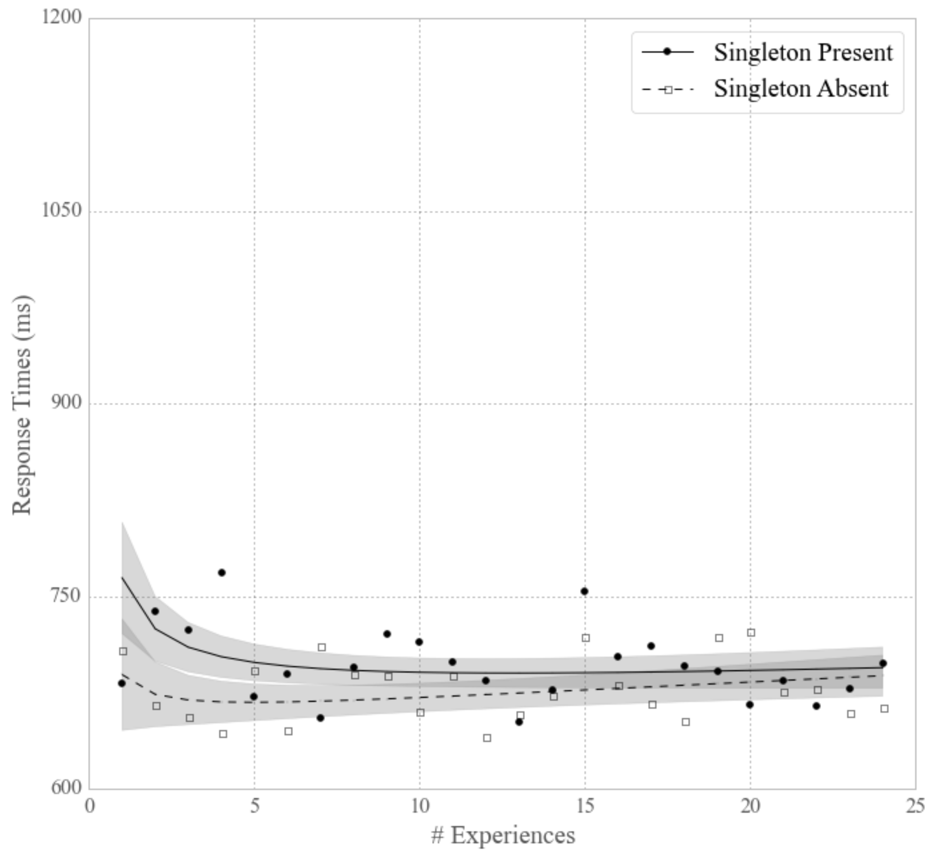

Figure 12. Experiment 2 RTs from the familiar singleton condition as a function of singleton presence and experience. The average RTs predicted by each observer's best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

These results demonstrate that when observers have previous experience with a particular color singleton, even if that experience occurred over 48 hours ago, these color singletons are not prioritized like completely novel color singletons. Instead, familiar color singletons distract observers less, and observers readily reject them. These results indicate that learned distractor rejection persists over extended breaks, demonstrating that learned distractor rejection relies on long-term memory.

Experiment 3: VWM's role in learned distractor rejection

Experiment 1 found that distractor experience is necessary for effective distractor rejection, but the exact memory system used to represent distractor experience remains uncertain. One possibility is the visual system stores *distractor rejection templates* in VWM (Woodman & Luck, 2007; Arita, Carlisle, & Woodman, 2012; Desimone & Duncan, 1995). According to Desimone and Duncan's Biased Competition theory of attention, observers use VWM to guide attention towards task relevant items. For example, when searching for an apple, an observer places a representation of an apple into VWM (i.e., target template) and this biases attention towards apple like items.

Arita and colleagues advanced the notion that VWM houses the templates that guide attention by positing that observers can also use VWM representations to bias attention away from irrelevant items (Arita et al., 2012; Woodman & Luck, 2007). In their experiment, Arita and colleagues presented observers with a cue that indicated what color the target would not be (i.e., *negative cue*). 6 items of one color appeared in the left hemifield and 6 items of another color appeared in the right hemifield. Observers identified the target faster when given a negative cue than when given a non-informative cue. The work of Arita et al suggests that effective distractor rejection relies on VWM resources.

It could be that observers learn to overcome distraction by loading a representation of the salient distractor into VWM, and use this representation to bias attention away from the salient distractor. Alternatively, Experiment 2's finding that learned distractor rejection persists for 48 hours, strongly suggests that LTM, rather than VWM, maintains learned distractor rejection. In fact, Beck and Hollingworth (2014) noted that because Arita and colleagues' (2012) used displays where color and hemifield were perfectly correlated, the experiment allowed participants to direct attention away from broad locations rather than specific features. Beck and Hollingworth found that observers exhibited equivalent performance when given a non-informative cue and when given a negative cue if there was no correlation between color and hemifield. Cunningham and Egeth (2014) found observers could exploit negative cues, if the same cue appeared on every trial and not when the cue changed every trial. These results suggest that the ability to use negative cues accrues gradually, rather than suddenly, as would be the case if VWM implemented negative cues. Based on the results of Experiment 2 and the lack of strong evidence for negative trial-by-trial cueing, it seems unlikely that VWM maintains learned distractor rejection.

While VWM does not seem to maintain representations that guide attention away from distractors, neural data supports the hypothesis that distractor rejection relies on working memory resources. For instance, Suzuki & Gottlieb (2013) found the dlPFC, an area involved in working memory (Curtis & D’Esposito, 2003), is necessary for effectively overcoming distraction. It is possible that the dlPFC implements distractor rejection instead of maintaining distractor rejection templates. If this is the case then preventing dlPFC resources from implementing distractor rejection, might interfere with learned distractor rejection. Specifically, while VWM might not maintain learned distractor rejection, VWM might be necessary for observers to consolidate learning. Without VWM resources, salient distractors might indefinitely distract observers and this is not because VWM maintains the learning, but because VWM is needed for observers to translate experience into learned distractor rejection.

Experiment 3 examines whether learned distractor rejection occurs without available VWM resources. If available VWM resources are necessary for observers to consolidate distractor experience, then learning will not occur when observers’ VWM resources are completed depleted. I will test this by filling VWM with extraneous items and examining whether observers still learn to overcome distraction. Observers will perform the same task as Experiment 1, but during the retention interval of a working memory task. Observers will always hold 4 items in VWM, completely depleting VWM resources (Luck & Vogel, 1997). If the consolidation of experience needed for learned distractor rejection relies on VWM, then depleting VWM resources with extraneous items will prevent observers from expressing learning. If the consolidation of experience used in learned distractor rejection does not rely on VWM resources, then filling VWM with extraneous items will not prevent observers from expressing learning. It could also be that filling VWM with extraneous items reduces the efficiency of encoding items into LTM (Atkinson & Shiffrin, 1968). In this case, learning will still occur, but slower than when observers perform the task without extraneous items in VWM.

Methods

Observers. Sixteen observers (10 females) from the University of Iowa psychology research participant pool participated for partial course credit. All observers reported normal vision and that they were not color blind. The Institutional Review Board approved this study and all observers gave informed consent. No observers from Experiment 3 had previously participated in Experiment 1.

Stimuli and Design. The methods of Experiment 2 follow the same methods as Experiment 1 except for one major difference (see Figure 13). In this experiment, observers completed a VWM task (Luck & Vogel, 1997) and performed the search task from Experiment 1 during the retention interval of this VWM task (Woodman, Vogel, & Luck, 2001). In the VWM task, observers began by receiving two numbers that served as an articulatory suppressant. Four colored squares then appeared evenly spaced around the circumference of an imaginary 4° circle.

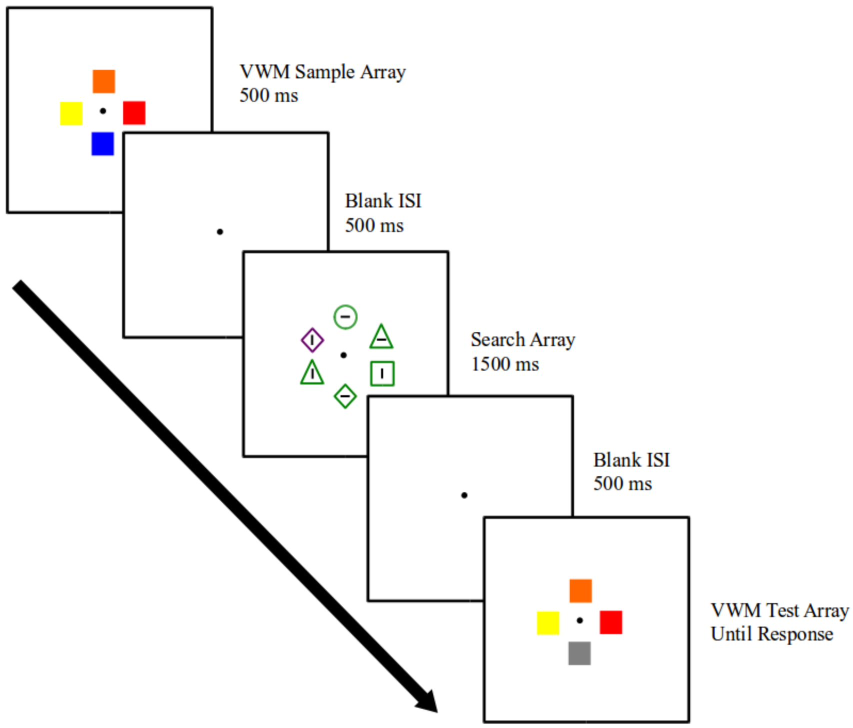

Figure 13. Sequence of events for Experiment 3. The VWM sample array appeared for 500 ms followed by a blank screen for 500 ms. After this, the search array appeared for 1500 ms or until response followed by another blank screen for 500 ms. Finally, the VWM test array appeared until response.

The colored squares appeared for 500 ms. The colored squares could be white (RGB 255,255,255), blue (RGB 20,20,204), brown (RGB 80,0,0), teal (0, 150, 150), grey (RGB 100, 100, 100), or tan (RGB 227, 168, 105). To prevent distraction by items matching the contents of working memory (Olivers et al., 2006), working memory items never matched any of the possible color singleton colors. 500 ms after the offset of the colored squares, the search array appeared and observers completed the same search task as Experiment 1. The search display appeared for 1500 ms or until response and was followed by a 500 ms blank ISI. After the blank ISI, observers saw a VWM probe array. On half the trials, all the items in the probe array matched the items from the sample array. On the other half, one of the items was a different color. Observers responded to whether they detected a change from the sample array via button press. Because of the additional keys necessitated by a second task, I changed key assignments. Observers used the “z,” “x,” “n,” and “m” keys. The “z” and “x” keys were used for one task and the “n” and “m” keys used for the other. Key assignment will be counterbalanced across participants.

Results and Discussion

RTs more than five standard deviations above the mean, faster than 300 ms, incorrect RTs, and RTs following an incorrect response were removed from the analysis. The RT trimming eliminated less than 0.2% of the data.

Observers successfully performed the visual working memory task, responding correctly in 67% of trials. Observers exhibited equivalent memory performance in singleton present and absent trials, $t(15) = 1.72$, $p > 0.10$.

The data and average predicted RTs by the best-fit power functions appear in Figure 14. The singleton present functions had an average RMSE of 104. The singleton absent function had an average RMSE of 87. At the beginning of the block, the singleton present functions predicted greater RTs than the singleton absent function, $t(15) = 3.71$, $p = 0.002$, demonstrating that color singletons initially distracted observers. β, the parameter representing the amount of learning between the beginning of the block and asymptote was greater in the singleton present than absent condition $t(15) = 3.31$, $p = 0.004$. α, the parameter representing learning rate, was significantly greater than 0 in the singleton present condition $t(15) = 3.84$, $p = 0.002$, but not in the singleton absent condition, $t(15) = 1.10$, $p > 0.28$. γ, the parameter representing asymptotic performance did not differ between the conditions, $t < 1$.

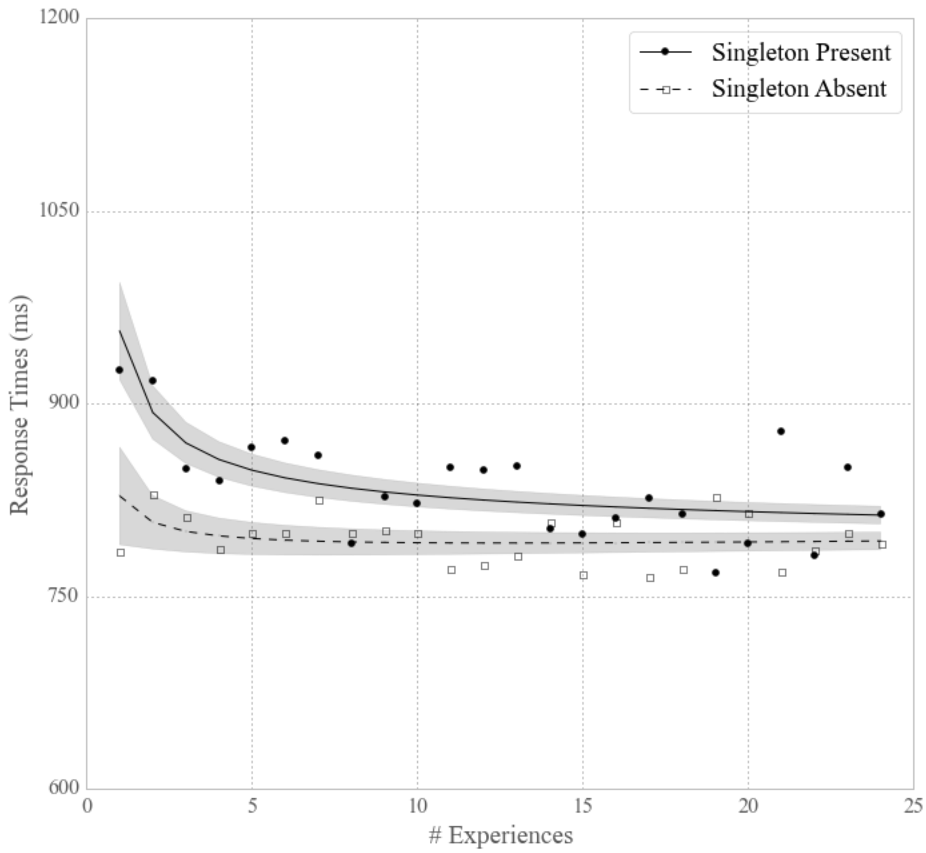

Figure 14. Experiment 3 RTs as a function of singleton presence and experience. The average RTs predicted by each observer's best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

These analyses demonstrate that RTs were initially much higher in the singleton present condition than the singleton absent condition and RTs decreased more in the singleton present condition than the singleton absent condition. Even though observers performed the search task with a full VWM load, observers still learned to better ignore distractors. This demonstrates that observers can consolidate distractor experience with a full VWM load. Interestingly, Figure 14 suggests that color singletons might have distracted observers at the end of blocks. Supporting this, the best fit power functions predict greater RTs in the singleton present condition at the end of the block, $t(15) = 3.22$, $p = 0.006$. Thus, while observers can still learn to ignore salient distractors, distraction persists throughout the block. Importantly, because the two functions have equivalent asymptotes, the best-fit power functions predict that performance will eventually converge in the two conditions. The VWM load just delays this convergence.

One potential problem with Experiment 3 is the VWM load might dilute the salience of color singletons. Specifically, maintaining four colored items in VWM might reduce the uniqueness (and thus saliency) of color singletons. While this might be the case, I still observed distraction at in Experiment 3, meaning that the color singleton was sufficiently salient.

These results demonstrate that filling VWM does not prevent observers from learning to better ignore salient color singletons. Nonetheless, filling VWM impede observers' ability to completely ignore color singletons as the color singleton still slowed RTs at the end of blocks. This difference between Experiments 1 and 3 strongly suggests that while visual working memory is not necessary for learned distractor rejection, it still plays a role in distractor rejection.

Experiment 4: Visually cueing the upcoming salient distractor's identity

When searching for a target, observers use VWM resources to bias attention towards relevant items (Desimone & Duncan, 1995; Woodman, Luck, & Schall, 2007). Cueing the identity of upcoming targets improves performance (Wolfe et al., 2004), and this cueing benefit seems to use VWM (Woodman, Vogel, & Luck, 2001; Woodman, Luck, & Schall, 2007; Carlisle et al., 2011). In fact, Cunningham & Egeth (2014) found observers could use cues to bias attention away from salient distractors. Interestingly, this only occurred if observers saw the same cue repeatedly and not if the cue changed on a trial-by-trial basis. This suggests that observers need experience with distractors before the distractors can be ignored (also see Lleras et al., 2008; Moher & Egeth, 2012; Tsal & Makovski, 2006). In Experiment 4, I will examine if observers can immediately reject salient distractors if they know the salient distractors identity before experiencing the salient distractor of if observers need experience rejecting salient distractors to effectively ignore them. If observers can explicitly bias attention away from salient distractor features then cueing the upcoming distractors' identity should remove any need to learn to ignore the salient distractor. Conversely, if learned distractor rejection relies on implicit experience-dependent memory systems, then cueing the upcoming distractors' identity will not prevent observers from needing to learn to ignore the salient distractors.

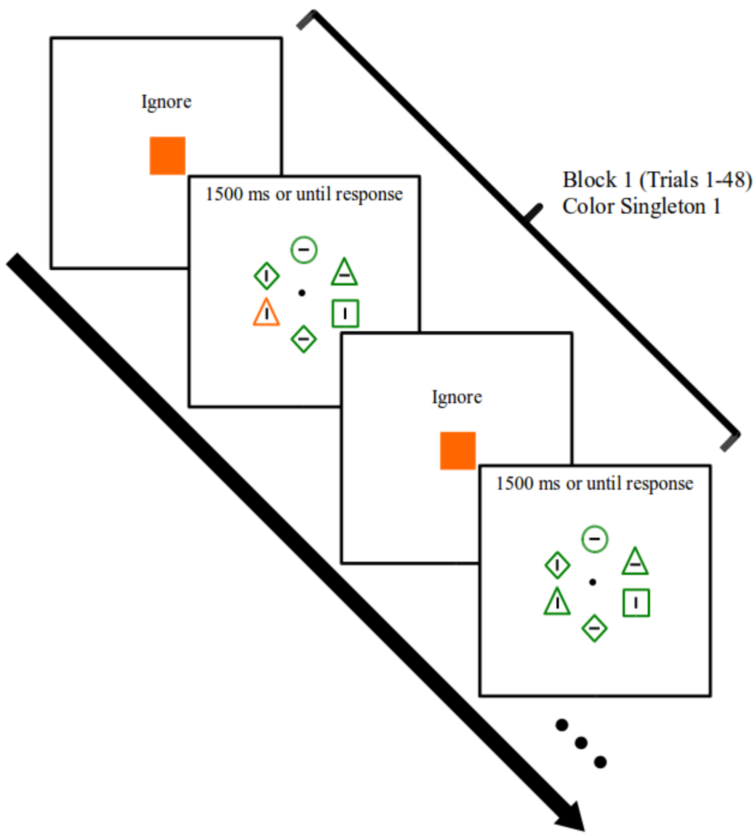

Figure 15. Sequence of events for Experiment 4. 1000 ms ITI and 250ms ISI not depicted. Before each trial, a prompt informed observers about the color of the salient distractor in that block.

Methods

Observers. Sixteen observers (11 females) from the University of Iowa psychology research participant pool participated for partial course credit. All observers reported normal vision and that they were not color blind. The Institutional Review Board approved this study and all observers gave informed consent. None of the observers in Experiment 4 participated in other experiments presented here.

Stimuli and Design. This experiment followed Experiment 1's methods except that observers were informed about the color of the upcoming salient distractor before each trial (see Figure 15). A prompt visually informed observers about the color singleton's color because pictorial cues are more effective than verbal cues (Wolfe et al., 2004). The prompt appeared for 250 ms. The prompt appeared before every trial. There was a 500 ms ISI between the prompt and the search array.

Results and Discussion

I used the same exclusion criteria as Experiment 1. The RT trimming excluded less than 0.1% of the data.

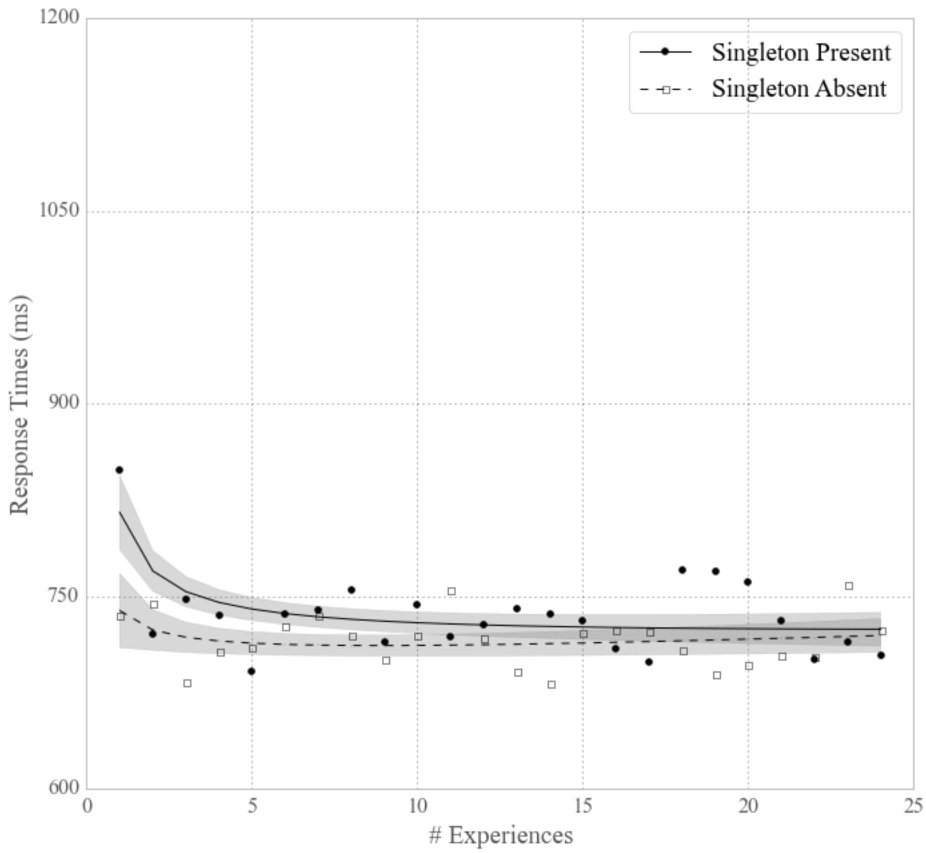

Figure 16. Experiment 4 RTs as a function of singleton presence and experience. The average RTs predicted by each observer's best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

The data and RTs predicted by average best-fit power functions appear in Figure 16. The singleton present functions had an average RMSE of 96. The singleton absent function had an average RMSE of 83. At the beginning of the block, the singleton present functions predicted greater RTs than the singleton absent function, $t(15) = 2.90$, $p = 0.01$, demonstrating that color singletons initially distracted observers. β, the parameter representing the amount of learning between the beginning of the block and asymptote, was greater in the singleton present condition, $t(15) = 3.26$, $p = 0.005$. α, the parameter representing learning rate, was not significantly greater than 0 in either the singleton present condition, $t(15) = 1.62$, $p > 0.12$, nor the singleton absent condition, $t < 1$. γ, the parameter representing initial performance, did not differ between the condition, $t < 1$. Observers initially responded slower when the color singleton was present than absent, indicating that the color singleton distracted participants. Performance improved in the singleton present condition, decreasing the amount of distraction exhibited. Interestingly, the learning rate did not differ from 0 in the singleton present condition. Nonetheless, observers exhibited a change in performance across the block; distraction early in the block and none late in the block.

When a colored square cues the color singleton, observers still must learn to ignore color singletons, as in Experiments 1, 2, and 3. It would not be fair to characterize the results of Experiment 1 and 3 as identical. In Experiment 4, the novel color singletons seem to produce less distraction than in Experiment 1. Supporting this observation, β was significantly greater in the color singleton present condition of Experiment 1 than the color singleton present condition of Experiment 4, $t(30) = 4.14$, $p < 0.001$. α of the singleton present condition did not differ between Experiment 4 and Experiment 1, $t(30) = 1.14$, $p > 0.26$. The salient distractor might have produced less distraction because of the short ISI between the cue and search task (500 ms). This short ITI means that the cue might have still been in iconic memory when the search array appeared (Coltheart, 1980), and this iconic memory trace may have prevented the salient distractor from being a true color singleton. Nonetheless, the fact that the salient distractor still initially distracted observers demonstrates that while cueing observers' with the color of upcoming salient distractors does reduce the color singleton's ability to distract observers, this cue does not obliterate distraction. This provides evidence that learned distractor rejection relies on an experience-dependent form of memory. Thus, salient distractors obligatorily initially distract observers, but after observers experience the salient distractors and learn to ignore them, observers can inhibit the features associated with this salient distractor (Moher et al., 2014).

Experiment 5: Verbally cueing the upcoming salient distractor's identity

Experiment 5 found that even when observers were cued with a salient distractor's identity, they still had to learn to ignore the item (see also Cunningham & Egeth, 2014). It is possible that the salient distractor initially distracted observers not because observers had to learn to ignore it but because observers automatically orient attention to items held in VWM (Olivers, Meijer, & Theeuwes, 2006). In Experiment 5, I verbally cued observers about the salient distractors identity. By verbally cueing observers, I still cue observers about the upcoming salient distractor, but this item will be verbally maintained. If automatic orienting to items in VWM created distraction in Experiment 4, then there will be no distraction in Experiment 5. If salient distractors distracted observers in Experiment 5 because observers must learn to ignore salient distractors, then Experiment 4 and 5 should produce similar results.

Methods

Observers: Sixteen observers (14 females) from the University of Iowa psychology research participant pool participated for partial course credit. All observers reported normal vision and that they were not color blind. The Institutional Review Board approved this study and all observers gave informed consent. None of the observers in Experiment 5 had participated in other experiments presented here.

Stimuli and Design. This experiment uses the same methods as Experiment 3, except observers were informed verbally, rather than visually, informed of the upcoming color singleton's color. For instance, the cue screen read, "ignore orange" when observers were in a block with orange color singletons.

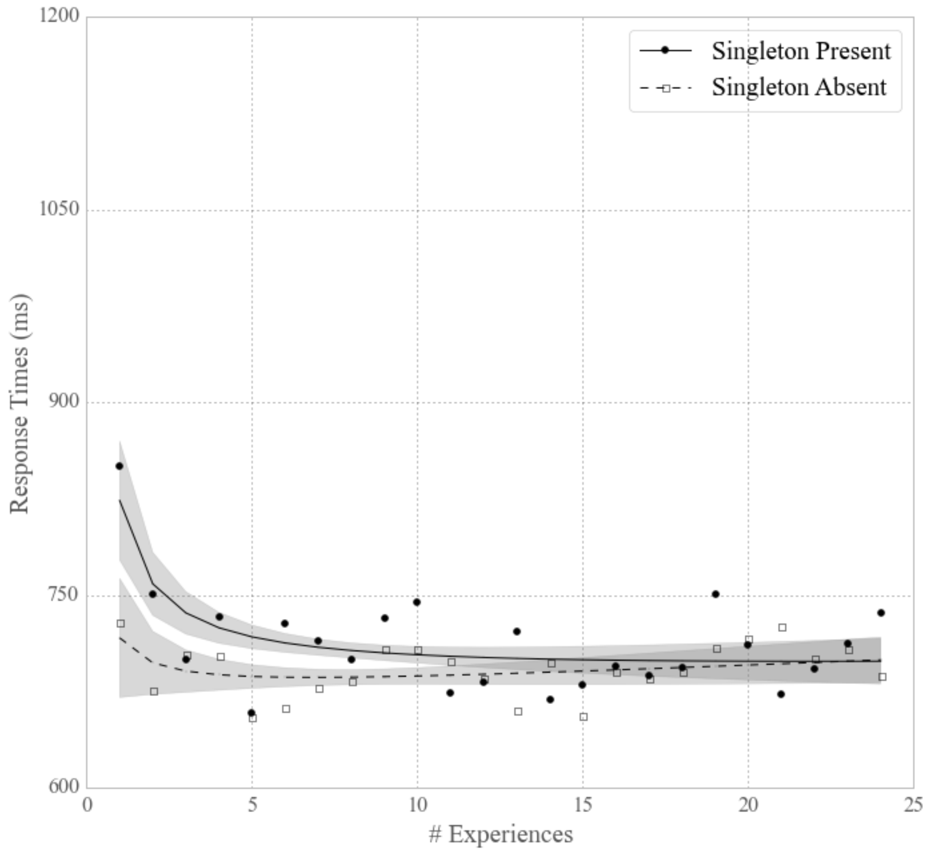

Figure 17. Experiment 5 RTs as a function of singleton presence and experience. The average RTs predicted by each observer's best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

Results and Discussion

I used the same exclusion criteria as Experiment 1. The RT trimming excluded less than 0.2% of the data.

The data and average RTs predicted by best-fit power functions appear in Figure 17. The singleton present functions had an average RMSE of 86. The singleton absent function had an average RMSE of 82. At the beginning of the block, the singleton present functions predicted greater RTs than the singleton absent function, $t(15) = 2.53$, $p < 0.05$, demonstrating that color singletons initially distracted observers. β, the parameter representing the amount of learning between the beginning of the block and asymptote, was marginally greater in the singleton present condition, $t(15) = 1.96$, $p = 0.06$. α was greater than zero in the singleton present condition, $t(15) = 2.24$, $p < 0.05$, but not in the singleton absent condition, $t < 1$, demonstrating that observers learned to ignore the color singleton in Experiment 5. γ, the parameter representing asymptotic performance did not differ between the two conditions, $t < 1$.

I compared the parameters between Experiments 4 and 5. I found no difference in either β, α, nor γ, all *t*s < 1. There were also no differences between the parameters representing the absent condition, all *t*s < 1.

In Experiment 5, I verbally cued observers about the identity of the upcoming distractor. This verbal cue should have created a verbal representation of the salient distractor, which is less likely to automatically guide attention (Olivers et al., 2006). Nonetheless, observers still learned to ignore the salient color singletons. These results replicate Cunningham and Egeth (2014) and demonstrate that even when observers have explicit knowledge about the upcoming distractors identity, they are unable to ignore salient distractors without previously experiencing them. The ability to ignore salient distractors accrues gradually with salient distractor experience, even if observers know the identity of the salient distractor. These results strongly suggest that learned distractor rejection relies on a form of implicit learning. Interestingly, Moher and colleagues (2011) found that observers could decrease distraction if they were told a distractor was likely on the upcoming trial, even if the probability of a distractor changed on a trial-by-trial basis. This result demonstrates that while observers cannot instantly implement a bias away from a specific distractor, they can broadly implement tighter attentional control. This suggests that preventing distraction consists of both feature specific inhibition (Moher et al., 2014) and a general tightening of attentional control (possibly by raising the selection threshold).

Experiment 6: Generalizability of learned distractor rejection

Experiment 2 demonstrates that learned distractor rejection relies on LTM, like skill learning. If learned distractor rejection relies on similar mechanisms to skill learning, it should exhibit similar learning patterns to skill learning. One prominent marker of long-term learning systems such as skill learning is environment heterogeneity's ability to influence the generalization of learning. Vatterott and Vecera (2012) found learned distractor rejection did not generalize across blocks. Instead, observers had to learn anew to reject novel color singletons in each block. Vatterott and Vecera used a very homogeneous learning environment though as each block contained a single identically colored color singleton. In Experiment 6, I ask if such a more heterogeneous stimulus environment can produce a generalizable distractor rejection strategy. Work in the skill learning literature demonstrates that heterogeneous stimulus environments help individuals generalize learning to novel environments (Schmidt & Bjork, 1992). For example, Catalano and Kleiner (1984) studied observers' ability to learn in a coincident-timing task. In this task, observers sat in front of a column of lights. The lights lit up one at time starting with the light most distant from the observer; each light lit in turn until the light closest to the observer was lit, creating the perception of an object moving toward the observer at a constant rate. The observers' task was to respond when the light closest to the observer lit. Observers could either practice this task in a training block where the timing between adjacent lights turning on was constant on each trial (the light moved at a constant speed towards observers) or observers could be in a training block where the light moved at a different speed each trial. All observers then transferred to a test block where they completed the same task, but the light moved either faster or slower than observers had experienced in the training block. Observers from the varied-speed practice group performed better in the test than the constant-speed practice group. This experiment demonstrates the importance of heterogeneous practice for generalizing skills and suggests that if learned distractor rejection uses the same learning principles as skill learning, a heterogeneous distractor environment will enable observers to reject novel distractors without previously experiencing these specific distractors.

Observers in Vatterott and Vecera's experiment experienced a single color-singleton in each block, providing observers with a homogeneous environment for learning to reject distractors. The current experiment tests whether a heterogeneous environment is better suited than a homogeneous environment for training observers to reject novel distractors. These experiments are largely modeled after Vatterott and Vecera (2012), with some key exceptions. In Experiment 6, half of the observers experienced a heterogeneous stimulus environment in which three different color-singletons appeared interspersed throughout three blocks (*Mixed group*). The other half of observers experienced a homogeneous stimulus environment in which a different singleton color appeared in each of three blocks (*Blocked group*), replicating Vatterott and Vecera (2012). Both groups were then tested with a novel distractor color in a final block of trials; this final block always contained a single distractor color, which appeared on half of the trials. If learned distractor rejection uses similar mechanisms as the skill learning described in Catalano and Kleiner (1984), then a novel color-singleton in the fourth block will initially interfere with the allocation of attention in the Blocked group, but not interfere with the allocation of attention in the Mixed group. If learned distractor rejection does not use the same learning principles as skill learning and is completely item specific, then the novel color-singleton will initially interfere with the allocation of attention in both groups.

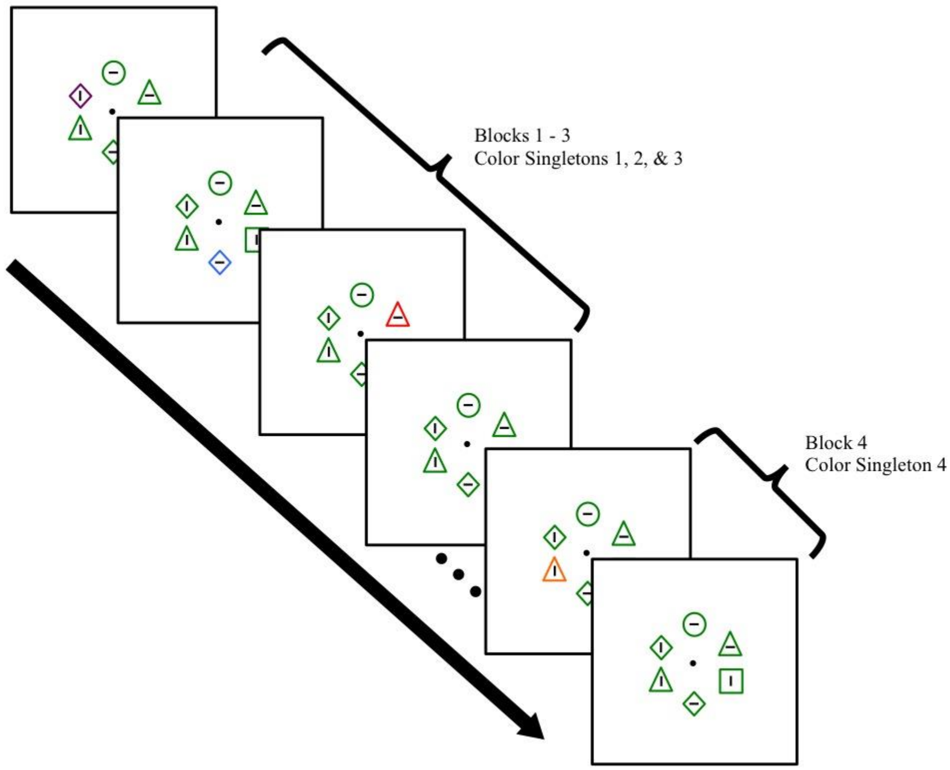

Figure 18. Sequence of events for mixed group in Experiment 6. 1000 ms ITIs not depicted. Color singletons one, two, and three all appeared intermixed in blocks one through 3 (color singletons still appear in 50% of trials). A single novel color singleton appears in block 4.

Methods

Observers. Thirty-two observers from the University of Iowa psychology research participant pool participated for partial course credit. All observers reported normal vision and that they were not color blind. The Institutional Review Board approved this study and all observers gave informed consent. None of the observers in Experiment 3 had participated in other experiments presented here. Sixteen observers (12 females) were in the Mixed group and sixteen observers (13 females) were in the Blocked group. Group assignment was decided pseudo-randomly.

Stimuli and Design. The methods are the same as Experiment 1 except for the following exceptions (see Figure 18). First, there will be two groups of observers. One group of observers (Blocked group) replicated Experiment 1, following the methods described in Experiment 1. The second group (Mixed group) experienced three different color singletons intermixed within the first three blocks. For example, an observer in the Blocked group might experience a red color singleton in block 1, a blue in block 2, and an orange in block 3. An observer in the Mixed group would experience red, blue, and orange color singletons all in the first block and then continue to experience these three color singletons in Blocks 2 and 3. Thus, all observers experience the same color singletons the same number of times. In the fourth block, all observers experience a single novel color singleton.

Results and Discussion

I used the same exclusion criteria as Experiment 1. This trimming eliminated less than 0.1% of the data.

I first estimated the power function parameters of data from blocks 1-3 independently for each observer in the Blocked group. The average RTs predicted by the best-fit power functions appear in Figure 19. The singleton present functions had an average RMSE of 183. The singleton absent function had an average RMSE of 169. At the beginning of the block, the singleton present functions predicted greater RTs than the singleton absent function, $t(15) = 2.96$, $p < 0.01$, demonstrating that color singletons initially distracted observers. β, the parameter representing the amount of learning between the beginning of the block and asymptote, was not different in the singleton present and absent conditions, $t(15) = 1.44$, $p > 0.17$. α, the parameter representing learning rate was greater than zero in the singleton present condition, $t(15) = 2.16$, $p < 0.05$, but not in the singleton absent condition, $t < 1$. γ, the parameter representing asymptotic performance was greater in the singleton present condition than the singleton absent condition, $t(15) = 2.93$, $p < 0.01$. These analyses roughly replicate Experiment 1, and indicate that when observers initially encountered a color singleton, it slowed response times, but observers gradually learned to ignore this color singleton, replicating Experiment 1.

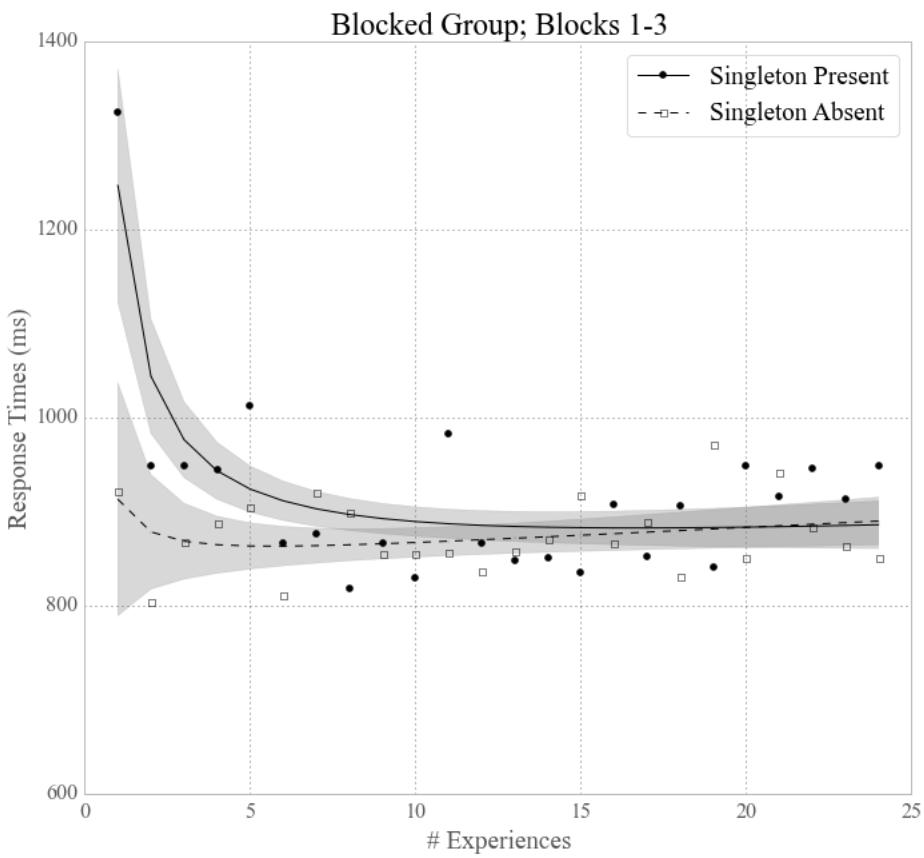

Figure 19. Blocked group, Block 1, Experiment 6 RTs as a function of singleton presence and experience. The average RTs predicted by each observer's best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

The intermixed color singletons prevented me from doing the same function fitting analysis with the Mixed group. Block wide analyses indicated that observers in the mixed group did learn to ignore the salient color singleton distractors (Vatterott, Mozer, & Vecera, *submitted*).

Next, I estimated the parameters of power functions fit to Block 4 of observers from both groups. The average RTs predicted by the best-fit power functions appear in Figure 20. The singleton present functions had an average RMSE of 347. The singleton absent function had an average RMSE of 276. At the beginning of the block, the singleton present functions predicted greater RTs than the singleton absent function, $t(15) = 2.77$, $p < 0.05$, demonstrating that color singletons initially distracted observers. β, the parameter representing the amount of learning between the beginning of the block and asymptote, did not differ between the singleton present and absent conditions, $t(15) = 1.74$, $p > 0.10$. The parameter representing learning rate, α, was significantly greater than zero in the singleton present condition, $t(15) = 3.16$, $p < 0.01$, but not singleton absent condition, $t(15) = 1.87$, $p > 0.08$. γ did not differ between the singleton present and absent conditions, $t < 1$. In the Mixed group, the singleton present functions had an average RMSE of 255 and the singleton absent functions had an average RMSE of 261. At the beginning of the block, the singleton present and absent functions predicted similar RTs, $t < 1$. β did not differ in the singleton present and absent conditions, $t < 1$. α was not significantly greater than zero in either the singleton present nor absent conditions, $t(15) = 1.48$, $p > 0.15$ and $t(15) = 1.05$, $p > 0.31$, respectively. γ did not differ between the singleton present and absent conditions, $t < 1$.

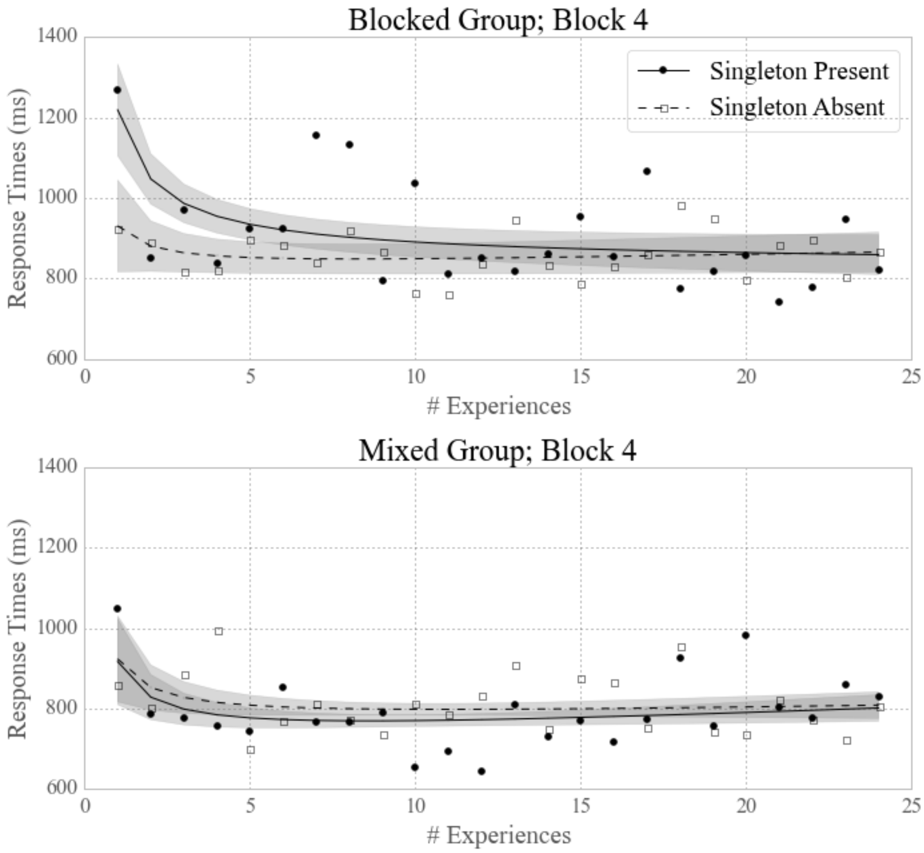

Figure 20. Block 4, Experiment 6 RTs as a function of singleton presence and experience. The average RTs predicted by each observer's best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

Comparing the two groups, β from the singleton present condition was not greater in the Blocked than Mixed group, $t(30) = 1.67$, $p = 0.10$. α from the singleton present condition was greater in the Blocked than Mixed group, $t(30) = 2.40$, $p < 0.05$. These analyses demonstrate that stimulus environment changes our ability to reject novel color singleton distractors.

Experiment 6 finds that experience with heterogeneous salient distractors enables observers to ignore novel salient distractors (the Mixed group), whereas experience with homogeneous salient distractors does not (the Blocked group). Thus, learned distractor rejection obeys at least one principle observed in skill learning: heterogeneous practice encourages generalization of learning (Schmidt & Bjork, 1992).

Two possible mechanisms could explain the generalization observed in this experiment. First, observers might have learned to discriminate the target and distractors more efficiently (Bacon & Egeth, 1994). For instance, it could be that observers learn to reject distractors not by learning to ignore a specific color value, but by narrowing their target template (Becker, Folk, & Remington, 2010); in my experiment, observers might begin by searching for a circle of any color and narrow their target template by experiencing salient color singletons. For example, experiencing a red color singleton might teach observers that the target is not red. The increased variability in the heterogeneous stimulus environment might encourage more drastic narrowing of the target template. Second, the heterogeneous stimulus environment might encourage observers to create more generalizable distractor rejection templates (Arita, Carlisle, & Woodman, 2012). For instance, observers might learn to reject specific distractor colors, but the heterogeneous stimulus environment causes observers to create less specific distractor rejection templates. These less precise distractor rejection templates make novel distractors more likely to fall within the color range specified by the rejection templates, and leave observers immune to interference by novel distractors.

In Experiment 6, the overall distractor statistics observed by Blocked and Mixed groups were identical: for both groups, over the first three blocks, the distractor color was chosen from three alternatives with equal probability. The statistics differ only in the *sequence* of colors. One explanation for why generalization is better with heterogeneity is that learning is recency based and performance in the final block of the experiment depends primarily on the statistics of the distractors in the penultimate block. If recency memory system, such as perceptual representation systems as in PoP (Tulving & Schacter, 1990), drives learning, one might expect it to be impossible for observers to learn a generalizable distractor rejection strategy in a blocked distractor environment. Conflicting with this account, past work from Vecera et al (2014) found that when observers learn to ignore a particular salient distractor, they maintain this learned distractor rejection template over the course of the entire experiment and Experiment 5 demonstrates that this learning exists for as long as 48 hours. If observers retain a learned distractor rejection template over the course of the experiment and these templates have any imprecision such that a green distractor rejection template leads to the rejection of all greenish distractors, then as observers gain experience with different distractors they will be more likely to immediately reject novel distractors. Additionally, from the perspective of skill acquisition, one would expect the learning I observe to be persistent and not purely recency based.

Experiment 7 was designed to resolve the apparent conflict between the result of Experiment 6, which suggests that learning is short lived, and results of Experiment 5, which suggests that learning persists for over 48 hours. Perhaps generalization may occur in principle with homogeneous environments, but there simply was not a sufficient number of homogeneous environments in Experiment 6 or a sufficient number of trials to observe generalization. To test this hypothesis, in Experiment 7, I presented a sequence of six, not four, homogeneous environments, and I extended the length of each block from 48 to 144 trials. Observers searched for a circle among heterogeneously shaped distractors over six blocks. A color-singleton appeared in half the trials, and the color of the color-singleton distractor was blocked. If learned distractor rejection is purely recency based, then increasing the number of trials and different color-singletons will not improve observers' ability to effectively ignore color-singleton distractors without previously experiencing them.

Experiment 7: Recent priming and learned distractor rejection

Experiment 7 investigated whether learned distractor rejection is purely recency based or if learned distractor rejection, like skill learning, persists across longer time scales. To distinguish between these hypotheses, I expanded the color-singleton distractor set from four colors to six and extended the length of each block to 144 trials. Observers began the experiment with a 60 trial training-block, as in Experiment 1, but by the end of the fourth block of trials, observers will have performed 636 trials searching for a constant target shape. By performing many trials, I hope to provide observers with a highly trained target template and robust distractor rejection templates, producing a strong test of distractor rejection templates' ability to persist across the experiment. If learned distractor rejection is purely recency based, then each time the salient color-singleton changes color, it should interfere with observers' allocation of attention, irrespective of the number of color-singletons previously experienced. If learned distractor rejection is robust and persistent, then color singletons will lose their ability to interfere with the allocation of attention as observers progress through the experiment, even after changes in the singleton's color.

Methods

Observers. Eighteen observers (12 females) from the University of Iowa psychology research participant pool participated for partial course credit. All observers reported normal vision and were not color blind. The Institutional Review Board approved this study and all observers gave informed consent. None of the observers in Experiment 3 had participated in other experiments presented here. More observers participated in Experiment 7 because I used six different salient distractor colors and needed more observers to fully counter-balance salient distractor order.

Stimuli and Design. Experiment 7 used the same methods as in Experiment 1 except for the following changes. Observers completed six blocks of trials instead of four, which means the observers experienced two additional color-singleton distractors (six total) and the test blocks were lengthened to from 48 to 144 trials. The two additional color-singleton colors were blue (RGB 0, 0, 255) and teal (RGB 0, 255, 255).

Results and Discussion

I used the same data trimming techniques as in Experiment 1. This trimming eliminated less than 0.1% of the data.

In Experiment 1, observers' RTs in the singleton present and absent conditions converged after about 15 trials in each condition (see Figure 7). Observers' performance changed as a function of experience in these initial trials, but after the two functions converge, performance did not change as a function of experience, and noise produced the remaining variability. Experiment 7 uses 144 trial blocks, increasing the percent of data representing trials after the singleton present and absent conditions converge and decreasing the percent of data representing pre-convergence performance. As one might expect, the initial trials of a block provide the most information for constraining a power function fit. To demonstrate this point, I conducted a simple simulation with noisy synthetic data generated from a power function. This simulation showed that modeling the data with more of the (noisy) tail led to poorer recovery of power-function parameters. The explanation for this phenomenon is that as the power function flattens, noise in the data overwhelms the signal. To combat this phenomenon, in Experiment 7, I fit power functions to only the first 24 singleton present and absent trials (the same number as Experiment 1).

Experiment 7 investigates whether learned distractor rejection is recency based or if learning persists across the experiment. To examine whether learning persists across the experiment, I fit separate power functions to observers' data from blocks 1-2, blocks 3-4, and blocks 5-6. The average best-fit power functions fit appear in Figure 21. The singleton present functions of blocks 1-2 had an average RMSE of 184. The singleton absent functions of blocks 1-2 had an average RMSE of 169. At the beginning of block 1-2, the singleton present functions predicted greater RTs than the singleton absent function, $t(17) = 2.15$, $p < 0.05$, demonstrating that color singletons initially distracted observers. The singleton present functions of blocks 3-4 had an average RMSE of 160. The singleton absent functions of blocks 3-4 had an average RMSE of 140. At the beginning of block 3-4, the singleton present functions predicted greater RTs than the singleton absent function, $t(17) = 2.23$, $p < 0.05$, demonstrating that color singletons initially distracted observers. The singleton present functions of blocks 5-6 had an average RMSE of 150. The singleton absent functions of blocks 5-6 had an average RMSE of 153. At the beginning of block 5-6, the singleton present functions predicted similar RTs as the singleton absent function, $t < 1$, demonstrating that color singletons were immediately ignored.

To investigate whether learning persists across the experiment I used the parameter representing the amount of learning, β, in a 3x2 repeated measures ANOVA with the factors Singleton Condition (Present; Absent) and Block Order (blocks1-2; blocks3-4; blocks5-6). The ANOVA found neither a main effect of Singleton Condition, $F(1,17) = 1.84$, $p > 0.19$, $\eta_p^2 = 0.10$. nor a main effect of Block Order, $F(2, 34) = 1.62$, $p > 0.21$, $\eta_p^2 = 0.17$. A marginal interaction subsumed both these main effects, $F(2,34) = 2.93$, $p = 0.07$, $\eta_p^2 = 0.29$, suggesting that β changed across singleton condition and Block. Planned comparisons found β was marginally greater in the singleton present than absent conditions of blocks 1-2 and blocks 3-4,

$t(17) = 2.02$, $p = 0.06$ and $t(17) = 1.75$, $p < 0.10$, respectively. β was not greater in the singleton present than absent condition of blocks 5-6, $t < 1$. These tests suggest that color singletons initially slowed RTs in blocks 1-2 and blocks 3-4, but not blocks 5-6.

The amount of learning changed between blocks as demonstrated by changes in the β parameter across blocks. β was greater in Blocks 1-2 than in Blocks 5-6, $t(17) = 4.07$, $p < 0.001$.

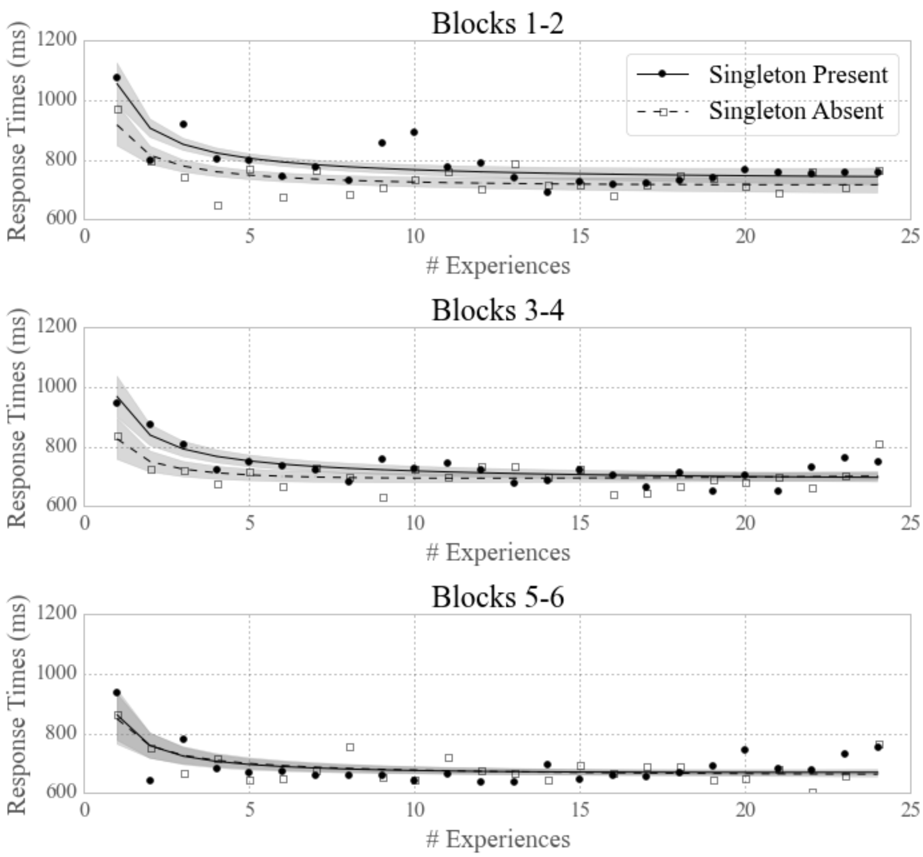

Figure 21. Experiment 7 RTs as a function of singleton presence and experience. The average RTs predicted by each observer's best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

Next, I investigated whether the parameter representing learning rate, α, varied across the experiment. α was greater than zero in the singleton present condition of blocks 1-2, $t(17) = 2.72$,

$p = 0.01$, blocks 3-4, $t(17) = 3.98$, $p < 0.001$, and marginally greater than zero in blocks 5-6, $t(17) = 1.89$, $p = 0.08$. α was also greater than zero in the singleton absent condition of blocks 1-2, $t(17) = 2.98$, $p < 0.01$ and blocks 5-6, $t(17) = 2.74$, $p = 0.01$. α was marginally greater than zero in blocks 3-4, $t(17) = 1.98$, $p > 0.06$. α values greater than zero in the singleton absent conditions indicate that observers improved performance even in trials without a color singleton.

γ, the parameter representing asymptotic performance did not differ between the singleton present and absent conditions of blocks 1-2, $t < 1$, nor blocks 3-4, $t(17) = 1.14$, $p > 0.26$. γ was greater in the singleton present trials of blocks 5-6 than the singleton absent trials of blocks 5-6, $t(17) = 2.62$, $p < 0.05$.

Experiment 7 found that when a novel color singleton was introduced at the beginning of blocks 1-4, attentional control failed to suppress the singleton. However, after about 15 experiences with the color singleton, observers were able to effectively reject the color singletons and color singletons ceased to interfere with the allocation of attention, replicating Experiment 1. In the initial trials of blocks 5 and 6, a novel color singleton did not interfere with the allocation of attention, in contrast to the initial trials of blocks 1-4. This finding suggests that learned distractor rejection, like skill learning, is not entirely recency based because even in a homogeneous stimulus environment, observers can learn to effectively reject salient color-singletons without previously experiencing these items. Thus, observers do retain some form of learning that persists over the time course of the experiment, indicating that learned distractor rejection operates over multiple time scales: the short time scale of learning to suppress a novel color singleton early in blocks 1-4, and the long time scale of learning to suppress a novel color singleton at the start of blocks 5 and 6.

Experiment 7 has a limitation, however, in that I cannot determine whether long time-scale learning arises from the greater number of distractor colors rejected or from the greater overall number of trials performed. That is, either performing more trials generally or experiencing more color-singleton distractor identities might have led to inter-block learning, which was expressed through the ability to immediately reject novel color singleton distractors in blocks 5 and 6. Experiment 8 sought to decouple these two factors by maintaining block length, but limiting the number of different color-singletons.

Experiment 8: Generalizability of learned distractor rejection in a homogeneous environment

Experiment 8 sought to examine what factor enabled observers in Experiment 8 to reject novel color-singleton distractors without previously experiencing them—whether it was the number of different distractor identities experienced or the overall number of trials. To distinguish between these two factors, observers in Experiment 8 performed the same experiment as Experiment 2, but the color singleton's color remained constant in blocks 1-4. Then observers experienced new color singletons in blocks 5 and 6. If performing more trials enabled observers to immediately reject salient color-singletons, then observers should immediately reject color singletons in blocks 5 and 6. If experiencing more color singletons enabled observers to immediately reject salient color-singletons, then color-singletons should initially interfere with the allocation of attention in blocks 5 and 6 because observers will only have experienced one color-singleton going into block 5 and two color-singletons going into block 6.

Methods

Participants. Eighteen observers (13 females) from the University of Iowa psychology research participant pool participated for partial course credit. All observers reported normal vision and were not color blind. The Institutional Review Board approved this study and all observers gave informed consent. None of the observers in Experiment 3 had participated in other experiments presented here.

Stimuli and Design. Experiment 8 used the same methods as Experiment 7 except that the color-singleton did not change color throughout blocks 1-4. Then, observers experienced new color-singletons in blocks 5 and 6. Thus, in total, observers experienced new color-singletons in blocks 1, 5, and 6. This design change resulted in only three different color-singleton colors. Experiment 3 used red, orange, and purple as the color-singleton colors.

Results and Discussion

I used the same data trimming techniques as Experiments 1. This trimming eliminated less than 0.1% of the data.

I used the same fitting procedure as Experiment 1. To identify whether an observer's ability to ignore novel color singletons depends on general task experience or the number of different color-singleton distractors previously experienced, I fit power functions to blocks 2-4 and blocks 5&6. The average best-fit power functions from blocks 2-4 and blocks 5&6 appear in the upper and lower panels of Figure 7, respectively.

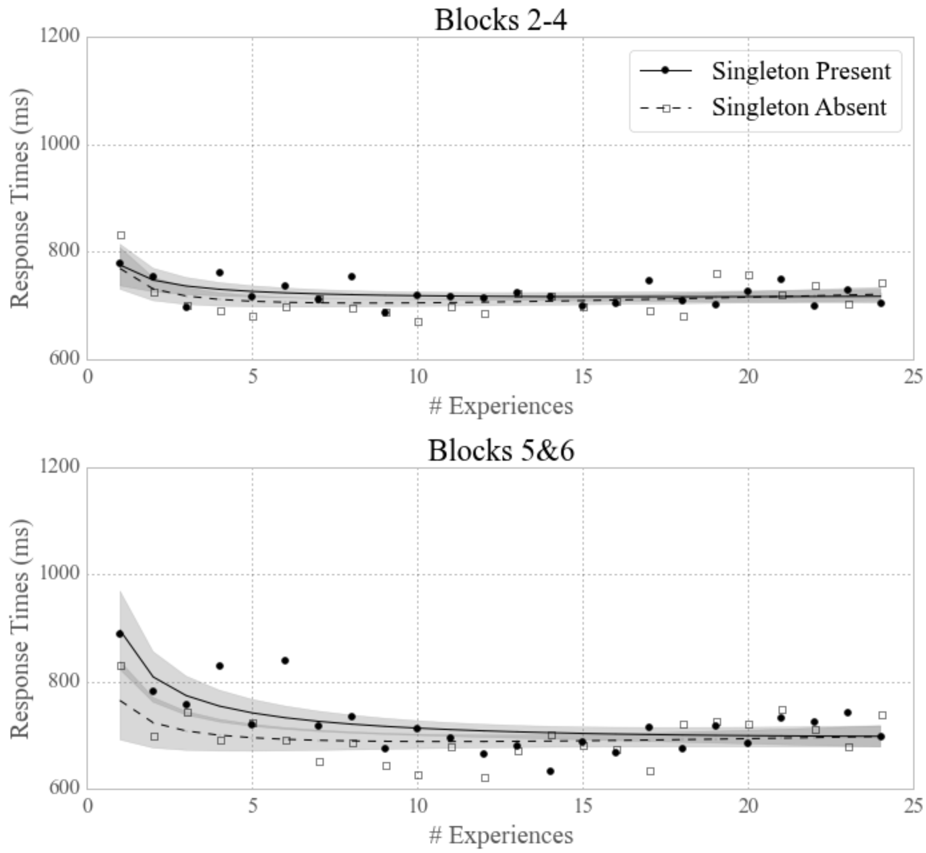

Figure 22. Experiment 8 RTs as a function of singleton presence and experience. The average RTs predicted by each observer's best-fit power function from singleton present and absent conditions appear as a solid and dotted line, respectively. Error bars depict the 95% confidence interval (Baguley, 2012; Loftus & Masson, 1994).

In blocks 2-4, the singleton present functions had an average RMSE of 86. The singleton absent function had an average RMSE of 95. At the beginning of the block, the singleton present and absent functions predicted similar RTs, $t < 1$, demonstrating that observers immediately ignored color singletons. β, the parameter representing the amount of learning between the beginning of the blocks and asymptotic performance was, did not differ in the singleton present than absent conditions, $t < 1$. In blocks 5-6, the singleton present functions had an average RMSE of 150. The singleton absent function had an average RMSE of 154. At the beginning of the block, the singleton present functions predicted marginally greater RTs than the singleton present functions, $t(15) = 1.94$, $p = 0.07$. γ did not differ between the singleton present and absent conditions, $t < 1$. Demonstrating that novel color singletons initially interfered with observers ability to allocate attention, in blocks 5&6, β was greater in the singleton present than absent condition, $t(17) = 2.15$, $p < 0.05$. In blocks 2-4, α, the parameter representing learning rate, was greater than zero in the singleton present condition, $t(17) = 2.66$, $p < 0.05$, but not in the singleton absent condition, $t < 1$. α was not greater than zero in the singleton present condition, $t(17) = 1.47$, $p > 0.15$. α was also not greater than zero in singleton absent condition of blocks 5&6, $t(17) = 1.21$, $p > 0.24$. γ did not differ between the singleton present and absent conditions of blocks 5-6, $t(17) = 1.82$, $p > 0.08$. Importantly, β was greater in the singleton present condition of blocks 5&6 than the singleton present condition of blocks 2-4, $t(17) = 3.21$, $p < 0.01$. This demonstrates that more learning occurred in blocks 5-6 than blocks 2-4.

I also fit a power function to data from blocks 1,5,&6. Adding Block 1 to the analyses did not change the basic pattern of results.

Experiment 8 demonstrated that observers could learn to effectively reject novel color-singletons even with previous homogeneous distractor experience. Experiment 8 sought to clarify whether additional task experience or experience with a greater number of color-singleton identities enabled observers to immediately reject novel color-singleton distractors. Experiment 8 found that novel color singletons in blocks 5&6 interfered with the allocation of attention, but familiar color singletons in blocks 2-4 did not. Thus, unlike Experiment 7 where color singletons did not initially interfere with the allocation of attention in blocks 5&6, color singletons in the same blocks did initially interfere with the allocation of attention in Experiment 8. Because the only difference between Experiment 7 and 8 is the number of color singletons previously experienced, the results indicate that experience with multiple color-singleton identities enabled observers in Experiment 7 to immediately reject novel color-singletons.

Before entering blocks 5 and 6, observers performed 60 practice trials and four 144 trial blocks (636 total trials). This large amount of task experience did not prevent novel color singletons from interfering with the allocation of attention. Although, extensive task experience might prevent novel color singletons from interfering with the allocation of attention, even without experiencing many distinct colors, Experiment 8 demonstrates that this experience would have to be extremely extensive.

CHAPTER 6. GENERAL DISCUSSION

Prominent models of attentional control (Bundesen, 1990; Desimone & Duncan, 1995) posit that a strong, specific bias towards targets prevents distraction. That is, if observers use a target template that completely segregates targets and salient distractors, attention will be biased towards targets, and this bias will prevent attention from visiting distractors. Thus, prominent models of attention predict that as long as salient distractors are sufficiently different from the target, their identity should not influence distraction. Experiment 1 tested this by periodically changing the identity of a salient distractor. Experiment 1 found that when the salient distractor changed identity, it initially distracted observers, but observers quickly learned to ignore the salient distractors. This result demonstrates that a specific bias towards target items is insufficient for preventing distraction. Instead, to effectively prevent distraction, observers must also learn to ignore salient distractors.

Distractor experience is critical for preventing distraction. Experiments 2 explored the memory system underlying this distractor experience. Experiment 2 evaluated whether LTM underlies learned distractor rejection. This experiment found that repeated, salient distractors experienced 48 hours later were less distracting than novel salient distractors. These results demonstrate that learned distractor rejection persists for over 48 hours, and strongly suggests that learned distractor rejection uses LTM. The finding that learned distractor rejection uses LTM falls in line with many other forms of experience-driven attentional control which also use LTM such as contextual cueing (Chun & Phelps, 1999), spatial probability cueing (Jiang et al., 2013), experience driven search strategies (Cosman & Vecera, 2013), and rewards influence on attentional control (Anderson & Yantis, 2013). These results conflict with demonstrations that distractor rejection relies on VWM (Arita et al., 2012). It is possible that VWM is involved in other parts of learned distractor rejection such as distractor disengagement (Fukuda & Vogel, 2011), but not involved in the learning portion of learned distractor rejection. Learned distractor rejection's reliance on LTM falls in line with models of attentional control, which posit that attentional control is learned via associating a particular image context with the target appearing at a particular location (Navalpakkam & Itti, 2005; Torralba et al., 2006). These models could easily account for learned distractor rejection by including a mechanism, which learns that particular features are never associated with the target.

Experiment 3 found that observers still learned to overcome distraction while holding a full VWM load. This results that VWM is not necessary for consolidating the experience used in learned distractor rejection. VWM is likely involved in distractor rejection. For instance, it could be that VWM resources help disengage from salient distractors (Fukuda & Vogel, 2011). Experiment 3 simply demonstrates that while VWM might be involved in distractor rejection, VWM resources are not necessary for learning to reduce distraction. Experiments 4 and 5 found that explicitly cueing observers with the identity of upcoming salient distractors did not obliterate distraction. These experiments demonstrate that learned distractor rejection cannot be manipulated by explicit processes. Instead, learned distractor rejection is an implicit, experience-dependent process. Distraction was notably reduced in Experiments 4-5 suggesting that explicit processes can reduce distraction (also see Moher et al., 2011). Importantly, experience with salient distractors is necessary to completely obliterate distraction.

Experiment 6 investigated the specificity of learned distractor rejection. This experiment found that while learned distractor rejection appears very specific Vatterott and Vecera (2012), observers can learn more general distractor rejection strategies if learning takes place in a more heterogeneous environment. This heterogeneous environment enables observers to learn a more generalizable distractor rejection strategy. Experiments 7 and 8 demonstrated that learned distractor rejection did not depend on perceptual representation systems (Tulving & Schacter, 1990). This, again, supports the idea that learned distractor rejection relies on LTM. Together, these results demonstrate that learned distractor rejection is a robust skill, which observers use to better control attention. This skill likely underlies performance benefits in complex tasks such as driving where observers must quickly learn what to attend to and what to ignore.

Models of Attentional Control and Learned Distractor Rejection

The results presented here could be implemented in prominent models of attentional control. TVA models attentional control as a race towards attentional selection (Bundesen, 1990; Bundesen et al., 2005; Bundesen et al., 2011). Items that win this race are selected and an item's importance, or similarity to important items, determines an item's likelihood to win the race. More important items have higher attentional weights, and are more likely to be selected. Bundesen (1990) states that an item's attentional weight is a non-negative number, meaning that item's cannot be explicitly devalued, but TVA could be implemented with all items possessing an attentional weight greater than zero and distractor items are devalued from this starting position. Thus, TVA has the structural components needed to implement the distractor devaluation observed here.

Bundesen (1990) describes the attentional weight of an item as being set by a homunculus. An important contribution of this work is to describe how we set attentional weights. Instead of relying on the idea of a homunculus for settings weights, this work demonstrates that experience dictates how we weight our attention towards important items and away from unimportant, salient items. An experience-dependent explanation does not need a homunculus, instead, when a feature is associated with the target, this feature receives a greater attentional weight than items not associated with the target. Importantly, as demonstrated in Experiment 1, unimportant, salient items are devalued. That is, when a conspicuous item is not the target, the attentional weight of this item decreases in an effort to prevent attention from visiting this item. Additionally, Experiment 1 demonstrates that even when searching for a specific target, the bias towards the target does not overwhelm the bias towards other items such that observers are immune to distraction. Instead, novel, salient distractors can distract observers, but observers can learn to prevent this distraction.

Experiments 2 investigated what memory system maintains learned distractor rejection. Experiment 2 found that observers maintained learned distractor rejection over extended delays, indicating that a LTM system maintains the attentional weights leading to learned distractor rejection. TVA offers no implemented memory system for maintaining attentional weights. Presumably, TVA could be outfitted with a LTM system for maintaining attentional weights. This memory system could store attentional weights in a persistent format and the relevant weights would be instantiated whenever a particular task demanded them. Experiment 3 demonstrated that VWM is not necessary for consolidating learned distractor rejection. That is, VWM resources are not necessary for learning to overcome distraction. TVA offers no mechanism for learning attentional weights, but TVA could easily be outfitted with a mechanism that tracks what features are associated with targets and what features are not. This mechanism would be independent of VWM and, therefore, would not interact with the amount of available VWM resources.

Biased Competition (Desimone & Duncan, 1995) could also model the results presented here. For instance, because Biased Competition posits that novel items are more salient than familiar items (Miller, Li, & Desimone, 1991), Biased Competition would predict that novel, salient distractors distract individuals. Individuals learn to reject novel distractors as the novel distractors become more familiar, and less salient. In this way, observers passively “learn” to reject salient distractors by becoming familiar with the distractors. Biased Competition does not explicitly address how individuals retain an item's familiarity, but Experiment 2 suggests that a LTM system stores this familiarity trace. Experiment 3 demonstrates that individuals can learn to ignore salient distractors without VWM resources. Biased competition would incorporate these results by implementing observers' ability to learn an item's familiarity independent of VWM. Biased competition can model the results presented here, but this model dictates that individuals would learn to ignore salient distractors not by explicitly devaluing them, but by becoming familiar with the items. As the items become less novel (i.e., more familiar) the items become less salient.

Experiments 4 and 5 investigate whether explicitly cueing an upcoming distractors identity prevents observers from needing to learn to reject salient distractors. Experiments 4 and 5 found that observers still needed to learn to reject salient distractors when they know upcoming distractor's identity. These experiments also investigate exactly how observers update their attentional control settings. Bundesen (1990) writes that a homunculus sets the attentional weights in TVA; this implies that explicit cognitive processes adjust attentional weights. For instance, when observers chooses to search for a green apple, TVA implements this by having the homunculus (i.e., experimenter) raise the attentional weight of green and round items. Experiments 4 and 5 demonstrate that observers cannot completely devalue salient distractors. Instead, observers need experience with a salient distractor to effectively reject them. Biased competition would model these results by implementing familiarity and explicit mental processes as independent. Biased competition would dictate that familiarity cannot be implemented without experience.

Experiment 6 found a heterogeneous learning environment encouraged broader generalization of learned distractor rejection. Again, because TVA offers no explicit mechanism for learning to devalue salient distractors, TVA would need an additional mechanism to model this finding. In order to assign an attentional weight to a given item, TVA must mark an item as a particular category. If this category is important, the item receives a greater attentional weight than if the assigned category is unimportant. A heterogeneous learning environment could encourage broader distractor categories than a homogeneous learning environment. The broader the distractor category, the more likely a novel distractor is to be categorized as an unimportant item. Alternatively, a heterogeneous learning environment could lead to observers devaluing distractor items less precisely. For instance, when an observer encounters a blue salient distractor, the heterogeneous learning environment could prevent the observer from devaluing a precise value blue, and, instead cause the observer to devalue all bluish items. These two implementations only differ in whether a heterogeneous distractor environment leads to a broad distractor category or less precise updating of attentional weights. Experiments 7 and 8 found that observers could learn to ignore salient distractors in a homogeneous learning environment. These experiments suggest that when observers experience a particular salient distractor, they do not update the attentional weighting of this precise item. Instead, there is some slop in what features are devalued. This causes observers to be more likely to reject a novel, salient distractor if they have previously experienced many salient distractors. Biased competition would implement these results by describing familiarity learning as malleable. In heterogeneous environments, observers are more conservative with what items are considered novel. That is, variability in salient distractors reduces the novelty of future salient distractors.

Together these results demonstrate that TVA could readily model many of the results presented here, but TVA would need additional mechanisms for updating attentional weights and for storing these attentional weights. The learning mechanism should be independent of VWM and should be less precise in variable environments. The memory mechanism that maintains attentional weights should persist over extended periods of disuse. Because of TVA's flexibility, TVA struggles to produce testable hypotheses involving learned distractor rejection. Instead, TVA's flexibility bends to produce almost any set of results. Biased competition, although similarly flexible, does produce a testable hypothesis. Biased competition models learned distractor rejection as a change in the familiarity of an item. This means that passive exposure to an item should eliminate any initial distraction by the item. Specifically, biased competition posits that if we become familiar with an item, through passive exposure, this item will be less distracting when it appears as a salient distractor. If learned distractor rejection involves devaluing items, as posited by TVA, then observers will still have to learn to ignore the now familiar salient distractor. Of course, if observers do not need to learn to ignore familiar salient distractors, then TVA could include novelty as a parameter affecting the attentional weighting of items, just as salience was added to TVA (Nordfang, Dyrholm, & Bundesen, 2013).

Stimulus-based and Goal-driven attentional control

Researchers have long debated the importance of stimulus-based and goal-driven attentional control. While I did not seek to resolve this issue, the results here do bear relevance to the debate. Specifically, my results add to the growing literature, which demonstrates that experience changes the relative importance of goal-driven and stimulus-based attentional control (Awh et al., 2012; Vecera et al., 2014; Wilder et al., 2011). The results presented here demonstrate that when observers first encounter an item, or task, observers exhibit more stimulus-driven control, and find salient items more distracting. After observers gain experience with the task, observers start to exhibit more goal-driven control. The results presented here demonstrate that this transfer between stimulus-driven and goal-driven attentional control occurs on a stimulus by stimulus basis, for instance, whenever a new stimulus appears, observers place more weight on stimulus-driven inputs.

The attentional control literature typically considers environment as the set of features within a static image (Theeuwes, 1992). For instance, saliency models typically determine saliency independent of past images (Itti & Koch, 2001), but the real world is not composed of static images. Instead, the world is a continuous flow, which encourages us to take into account the recent past when calculating salience (Baldi & Itti, 2010). This posits that experiencing many bright orange items in the recent past makes current bright orange items less salient. In other words, novels items are more salient than familiar items. Miller, Li, and Desimone (1991) found evidence for a novelty bias with monkeys performing a delayed match to sample task. In this task, monkeys first viewed a stimulus, which they were to remember then a number of test stimuli. Monkeys had to indicate when a test stimulus matches the memory stimulus. Miller and colleagues found cells responded more vigorously when the cells preferred stimulus was novel than when it was familiar. Importantly, these cells maintained selectivity in that they did not respond to all novel stimuli, only the novel stimuli to which the cell was tuned. This provides a passive mechanism by which the visual system can preferentially attend to novel, salient, information. Importantly, this mechanism, posited by Biased Competition (Desimone & Duncan, 1995), does not invoke any goal-driven changes in the visual system.

Limitations

One limitation in this series of experiments is I only demonstrated learned distractor rejection in a single attentional control paradigm- a variant of the additional singleton paradigm

(Bacon & Egeth, 1994; Theeuwes, 1992). The mechanism underlying distraction costs in the additional singleton paradigm could arise from either attentional capture by the salient color singleton (Theeuwes, 1992) or a filtering cost of the salient color singleton (Folk & Remington, 1998). I remained agnostic as to the exact mechanism underlying the distraction costs observed here. Whether attentional capture or a filtering cost creates distraction, distraction costs do reflect a reduced efficiency in allocating attention to the target. In other words, whatever the mechanism underlying the distraction costs observed here, there is still an attentional control cost. Using a single paradigm in these experiments also questions the generalizability of the current findings. Nonetheless, Vecera et al (2014) found similar experience driven effects using a cueing paradigm (Folk et al., 1992), which suggests that the same learning mechanisms observed here also exist in other paradigms.

One significant limitation of the work presented here is the relatively homogeneous participant population. Our observers were all University of Iowa students, and the University of Iowa has a particularly homogeneous student population. Only 15% of University of Iowa students are from a minority background. It is possible that the homogeneous sample used here produced a set of results that will not generalize to other populations (Henrich, Heine, & Norenzayan, 2010). Alternatively, while there are individual differences in distractibility (Fukuda & Vogel, 2009; Ophir, Nass, & Wagner, 2009), distraaction almost certainly occurs in every human (and many non-human) population, so learning to overcome distraction also likely occurs across the human population. It is possible that different populations use different strategies to overcome distraction. Giving the experiments presented here to different populations across the globe could test this interesting possibility.

Conclusion

Experience with salient distractors helps prevents salient distractors from distracting observers. This simple result demonstrates that efficient attentional control depends not only on a specific target template, but also depends on experience with salient distractors. Like other forms of experience driven attentional control, learned distractor rejection relies on LTM. This finding is fitting as LTM has the capacity and longevity to effectively guide attention away from a wide variety of salient distractors in the present and the distant future. This demonstrates that we should not view attentional control as simply using stimulus-driven and goal-drive inputs. Instead, attentional control dynamically changes based on experience with particular items within a task. Thus, the ability to control attention is a skill that is gradually honed across experience. With little experience, this skill is relatively simple and relies on the basic heuristic of saliency, but as observers gain more experience with the task, this skill becomes more honed and adapts to a tasks particular nuances, leading to effective, goal-driven attentional control.

REFERENCES

Allman, J., Miezen, F., & McGuinness, E. (1985). Stimulus specific response from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparison in visual neurons. *Annual Reviews of Neuroscience, 8*(1), 407-430. doi: 10.1146/annurev.ne.08.030185.002203

Anderson, B.A., Laurent, P.A., Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*(25), 10367-10371. doi: 10.1073/pnas.1104047108

Anderson, B.A., Yantis, S. (2013). Persistence of value-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 6-9. doi: 10.1037/a0030860

Annac, E., Manginelli, A.A., Pollman, S., Shi, Z., Müller, H.J., & Geyer, T. (2013). Memory under pressure: Secondary-task effects on contextual cueing of visual search. *Journal of Vision, 13*(13):6, 1-15. doi: 10.1167/13.13.6

Arita, J.T., Carlisle, N.B., & Woodman, G.F. (2012). Templates for rejection: Configuring attention to ignore task-irrelevant features. *Journal of Experimental Psychology: Human Perception and Performance, 38*(3), 580-584. doi: 10.1037/a0027885

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation, 2,* 89-195

Awh, E., Belopolsky, A.V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*(8), 437-443. doi: 10.1016/j.tics.2012.06.010

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics, 55*(5), 485-96. doi: 10.3758/BF03205306

Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behavior Research Methods*, *44*(1), 158-175. doi: 10.3758/s13428-011-0123-7

Baldi, P., & Itti, L. (2010). Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Networks*, *23*(5), 649-666. doi: 10.1016/j.neunet.2009.12.007

Beck, V., & Hollingworth, A. (2014). Can attention be guided efficiently by a negative template? *Journal of Vision, 14*(10), 1050

Becker, S.I. (2008). The stage of priming: Are intertrial repetition effects attentional or decisional? *Vision Research*, *48*(5), 664-684. doi: 10.1016/j.visres.2007.10.025

Becker, S.I., Folk, C.L., & Remington, R.W. (2010). The role of relational information in contingent capture. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1460-1476. doi: 10.1037/a0020370

Brady, T.F., Konkle, T., Alvarez, G.A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105*(38), 14325-14329. doi: 10.1073/pnas.0803390105

Brady, T.F., & Chun, M.M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance, 33*(4), 798-815. doi: 10.1037/0096-1523.33.4.798

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433-464

Brooks, D.I., Rasmussen, I.P., & Hollingworth, A. (2010). The nesting of search contexts within natural scenes: Evidence from contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1406-1418. doi: 10.1037/a0019257

Bruce, N.D.B., & Tsotsos, J.K. (2009). Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision, 9*(3), 5. doi: 10.1167/9.3.5

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*(4), 523-547. doi: 10.1037/0033-295X.97.4.523

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review, 112*(2), 291-328. doi: 10.1037/0033-295X.112.2.291

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2011). A neural theory of visual attention and short-term memory (NTVA). *Neuropsychologia, 49*(6), 1446-1457. doi: 10.1016/j.neuropsychologica.2010.12.006

Carlisle, N.B., Arita, J.T., Pardo, D., Woodman, G.F. (2011). Attentional templates in visual working memory. *Journal of Neuroscience, 31*(25), 9315-9322. doi: 10.1523/JNEUROSCI.1097-11.2011

Catalano, J.F., & Kleiner, B.M. (1984). Distant transfer in coincident timing as a function of variability of practice. *Perceptual and Motor Skills, 58*(3), 851-856.

Chelazzi, L., Miller, E.K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in the inferior temporal cortex. *Nature, 363*, 345-347. doi: 10.1038/363345a0

Cherry, E.C. (1953). Some experiments on the recognition of speech, with One and with Two Ears. *The Journal of the Acoustical Society of America, 25*(5), 975-979. doi: 10.1121/1.1907229

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*(1), 28-71. doi: 10.1006/cogp.1998.0681

Chun, M.M., & Jiang, Y. (2003). Implicit, Long-Term Spatial Contextual Memory. *Journal of Experimental Psychology: Learning, Memory, and Performance, 29*(2), 224-234. doi: 10.1037/0278-7393.29.2.224

Chun, M.M., & Phelps, E.A. (1999). Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. *Nature Neuroscience, 2*(9), 844-847. doi: 10.1038/12222

Cohen, N.J., & Eichenbaum, H. (1993). *Memory, Amnesia, and the Hippocampal System.* Cambridge: MIT Press.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics, 27*(3), 183-228. doi: 10.3758/BF03204258

Cosman, J.D., & Vecera, S.P. (2012). Context-dependent control over attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 39*(3), 836-848. doi: 10.1037/a0030027

Cosman, J.D., & Vecera, S.P. (2013). Learned control over distraction is disrupted in amnesia. *Psychological Science, 24*(8), 1585-1590. doi: 10.1177/0956797613475632

Cosman, J.D., & Vecera, S.P. (2014). Establishment of an attentional set via statistical learning. *Journal of Experimental Psychology: Human Perception and Performance, 40*(1), 1-6. doi: 10.1037/a0034489

Crossman, E.R.F.W. (1959). A theory of the acquisition of speed-skill. *Ergonomics, 2*(2), 153-166. doi: 10.1080/00140135908930419

Cunningham, C.A., & Egeth, H.E. (2014). When constant rejection is a good thing: The effects of learning on efficient feature inhibition. *Object Perception, Attention, and Memory (OPAM),* Long Beach, California.

Curtis, C.E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences, 7*(9), 415-423. doi: 10.1016/S1364-6613(03)00197-9

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Reviews of Neuroscience, 18*(1), 193-222. doi: 10.1146/annurev.ne.18.0030195.001205

Downing, P.E. (2000). Interactions between visual working memory and selective attention. *Psychological Science, 11*(6), 467-473. doi: 10.1111/09280.00290

Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433-458. doi: 10.1037/0033-295X.96.3.433

Egeth, H.E., Leonard, C.J., & Leber, A.B. (2010). Why salience is not enough: Reflections on top-down selection in vision. *Acta Psychological, 135*(2), 130-132. doi: 10.1016/j.actpsy.2010.05.012

Egeth, H.E., Virzi, R.A., & Gabart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance, 10*(1), 32-39. doi: 10.1037/0096-1523.10.1.32

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1030-44. doi: 10.1037/0096-1523.18.4.1030

Folk, C.L., & Remington, R.W., & Wright, J.H. (1994). The structure of attentional control: contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance, 20*(2), 317-329. doi: 10.1037/0096-1523.20.2.317

Folk, C. L., Remington, R. W. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 847-858. doi: 10.1037/0096-1523.24.3.847

Fukuda, K., & Vogel, E.K. (2009). Human variation in overriding attentional capture. *The Journal of Neuroscience, 29*(27), 8726-8733. doi: 10.1523/JNEUROSCI.2145-09.2009

Fukuda, K., & Vogel, E.K. (2011). Individual differences in recovery time from attentional capture. *Psychological Science, 22*(3), 361-368. doi: 10.1177/0956797611398493

Geng, J.J. (2014). Attentional mechanisms of distractor suppression. *Current Directions in Psychological Science, 23*(2), 147-153. doi: 10.1177/0963721414525780

Geng, J.J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics, 67*(7), 1252-1268. doi: 10.3758/BF03193557

Han, S.W., & Kim, M.S. (2009). Do the contents of working memory capture attention? Yes, but cognitive control matters. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1292-1302. doi: 10.1037/a0016452

Heathcote, A., Brown, S., & Mewhort, D.J.K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*(2), 185-207. doi: 10.3758/BF03212979

Henrich, J., Heine, S.J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 81-83. doi: 10.1017/s0140525X0999152X

Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering, 9*(3), 90-95. doi: 10.1109/MCSE.2007.55

Itti, L. & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194-203. doi: 10.1038/35058500

Jiang, Y.V., Swallow, K.M., Rosenbaum, G.M., & Herzig, C. (2013). Rapid acquisition but slow extinction of an attentional bias in space. *Journal of Experimental Psychology: Human Perception and Performance, 39*(1), 87-99. doi: 10.1037/a0027611

Kawahara, J.I. (2010). Identifying a "default" visual search mode with operant conditioning. *Acta Psychologica, 135*(1), 38-49. doi: 10.1016/j.actpsy.2010.05.002

Kristjánsson, Á., Vuilleumier, P., Schwartz, S., Macaluso, E., & Driver, J. (2007). Neural basis for priming of pop-out during visual search revealed with fMRI. *Cerebral Cortex, 17*(7), 1612-1624. doi: 10.1093/cercor/bhl072

Leber, A. B., & Egeth, H. E. (2006). It's under control: top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, *13*(1), 132-8. doi: 10.3758/BF03193824

Leber, A. B., Kawahara, J.I., & Gabari, Y. (2009). Long-term abstract learning of attentional set. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1385-97. doi: 10.1037/a0016470

Lee, H., Mozer, M.C., & Vecera, S.P. (2009). Mechanisms of priming of pop-out: Stored representations or feature-gain modulations? *Attention, Perception, & Psychophysics, 71*(5), 1059-1071. doi: 10.3758/APP.71.5.1059

Lleras, A., Kawahara, J.I., Wan, X., & Ariga, A. (2008). Intertrial inhibition of focused attention in pop-out search. *Perception & Psychophysics*, *70*(1), 114-131. doi: 10.3758.PP.70.1.114

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*(4), 476–490. doi: 10.3758/BF03210951

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492-527. doi: 10.1037/0033-295X.95.4.492

Logan, G.D. (2002). An instance theory of attention and memory. *Psychological Review, 109*(2), 376-400. doi: 10.1037/0033-295X.109.2.376

Luck, S.J. (2008). Visual short-term memory. In S.J. Luck & A. Hollingworth (Eds.), *Visual Memory* (pp. 43-85). New York: Oxford University Press.

Luck, S.J., Chelazzi, L., Hillyard, S.A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology, 77*(1), 24-42.

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279-281. doi: 10.1038/36846

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22*(6), 657-72. doi: 10.3758/BF03209251

Manginelli, A.A., Langer, N., Klose, D., & Pollmann, S. (2013). Contextual cueing under working memory load: Selective interference of visuospatial load with expression of learning. *Attention Perception & Psychophsics, 75*(6), 1103-1117. doi: 10.3758/s13414-013-0466-5

McKinney, W. (2010). Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference,* 51-56.

Miller, E.K., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science, 263*((5146), 520-522. doi: 10.1126/science.8290960

Miller, E.K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science, 254*(5036), 1377-1379. doi: 10.1126/science.1962197

Moher, J., Abrams, J., Egeth, H.E., Yantis, S., & Stuphorn, V. (2011). Trial-by-trial adjustments of top-down set modulate oculomotor capture. *Psychonomic Bulletin & Review, 18*(5), 897-903. doi: 10.1037/s13423-011-0118-5

Moher, J., & Egeth, H.E. (2012). The ignoring paradox: Cueing distractor features leads first to selection, then to inhibition of to-be-ignored items. *Attention, Perception, & Psychophysics, 74*(8), 1590-1605. doi: 10.3758/s13414-012-0358-0

Moher, J., Lakshmanan, B.M., Egeth, H.E., & Ewen, J.B. (2014). Inhibition drives early feature-based attention. *Psychological Science, 25*(2), 315-324. doi: 10.1177/0956797613511257

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research, 45*(2), 205-231. doi: 10.1016/j.visres.2004.07.042

Newell, A., & Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (Ed.), *Cognitive Skills and their Acquisition* (pp.1-55), Hillsdale, NJ; Erlbaum.

Nordfang, M., Dyrholm, M., & Bundesen, C. (2013). Identifying bottom-up and top-down components of attentional weight by experimental analysis and computational modeling. *Journal of Experimental Psychology: General, 142*(2), 510-535. doi: 10.1037/a0029631

Nosofsky, R.M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(4), 700-708. doi: 10.1037/0278-7393.14.4.700

Nosofsky, R.M., & Palmeri, T.J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104*(2), 266-300. doi: 10.1037/0033-295X.104.2.266

Oliphant, T.E. (2007). Python for scientific computing. *Computing in Science & Engineering, 9*(3), 10-20. doi: 10.1109/MCSE.2007.58

Olivers, C.N.L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1243-1265. doi: 10.1037/0096-1523.32.5.1243

Olson, I.R., & Chun, M.M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition, 9*(3), 273-302. doi: 10.1080/13506280042000162

Ophir, E., Nass, C., & Wagner, A.D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences, 106*(37), 15583-15587. doi: 10.1073/pnas.0903620106

Pashler, H. (1988). Cross-dimensional interaction and texture segregation. *Perception & Psychophysics*, *43*(4), 307-18. doi: 10.3758/BF03208800

Pérez, F., & Granger, B.E. (2007). IPython: A system for interactive scientific computing. *Computing in Science & Engineering, 9*(3), 21-29. doi: 10.1109/MCSE.2007.53

Phillips, W.A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*(2), 283-290. doi: 10.3758/BF03203943

Posner, M.I., Snyder, C.R., & Davidson, B.J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*(2), 160-174. doi: 10.1037/0096-3445.109.160

Postle, B. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139*(1), 23-38. doi: 10.1016/j.neuroscience.2005.06.005

Raymond, J.E., & O'Brien, J.L. (2009). Selective visual attention and motivation: The consequences of value learning in an attentional blink task. *Psychological Science, 20*(8), 981-988. doi: 10.1111/j.1467-9280.2009.02391.x

Rosenholtz, R. (1999). A simple saliency model predicts a number of motion popout phenomena, *Vision Research, 39*(19), 3157-3163. doi: 10.1016/S0042-6989(99)00077-2

Rosenholtz, R., Li, Y., Nakano, L. (2007). Measuring visual clutter. *Journal of Vision, 7*(2):17, 1-22. doi: 10.1167/7.2.17

Salient (n.d.). In *Merriam-Webster Online*. Retrieved from http://www.merriam-webster.com/dictionary/salient.

Schacter, D.L., & Buckner, R.L. (1998). Priming and the brain. *Neuron, 20*(2), 185-195. doi: 10.1016/S0896-6273(00)80448-1

Schmidt, R.A., & Bjork, R.A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207-217. doi: 10.1111/j.1467-9280.1992.tb00029.x

Schneider, W., & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1-66. doi: 10.1037/0033-295X.84.1.1

Shaw, M.L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance, 3*(2), 201-211. doi: 10.1037/0096-1523

Shepard, R. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika, 22*(4), 325- 345. doi: 10.1007/BF02288967

Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*(2), 127-190. doi: 10.1037/0033-295X.84.2.127

Soto, D., Heinke, D., Humphreys, G.W., & Blanco, M.J. (2005). Early, involuntary, top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance, 31*(2), 248-261. doi: 10.1037/0096-1523.31.2.248

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monograpghs: General and applied, 74*(11), 1-29. doi: 10.1037/h0093759

Suzuki, M., & Gottlieb, J. (2013). Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. *Nature Neuroscience, 16*(1), 98-104. doi: 10.1038/nn.3282

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51*(6), 599-606. doi: 10.3758/BF03211656

Theeuwes, J. (2010). Top-down and bottom-up selection of visual selection. *Acta Psychologica, 135*(2), 77-99. doi: 10.1016/j.actpsy.2010.02.006

Torralba, A., Oliva, A., Castelhano, M.S., & Henderson, J.M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766-786. doi: 10.1037/0033-295X.113.4.766

Tulving, E., & Schacter, D.L. (1990). Priming and human memory systems. *Science, 247*(4940), 301-306. doi: 10.1126/science.2296719

Travis, S.L., Mattingley, J.B., & Dux, P.E. (2013). On the role of working memory in spatial contextual cueing. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 39*(1), 208-219. doi: 10.1037/a0028644

Treisman, A.M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97-136. doi: 10.1016/0010-0285(80)90005-5

Tsal, Y., & Makovski, T. (2006). The attentional white bear phenomenon: The mandatory allocation of attention to expected distractor locations. *Journal of Experimental Psychology: Human Perception and Performance, 32*(2), 351-363. doi: 10.1037/0096-1523.2.351

van der Walt, S., Colbert, C., & Varoquaux, G. (2011). The Numpy array: A structure for efficient numerical computation. *Computing in Science & Engineering, 13*(2), 22-30. doi: 10.1109/MCSE.2011.37

Vatterott, D.B., Hout, M.C., Goldinger, S.D., & Vecera, S.P. (*under revision*). Experience modulates attention to irrelevant distractors: Evidence from saccade deviation.

Vatterott, D.B., Mordkoff, J.T., & Vecera, S.P. (*in prep*). Adaptive filtering explains color singleton related slowing in the additional singleton paradigm.

Vatterott, D.B., Mozer, M.C., & Vecera, S.P. (*in prep*). Learning distractor rejection is a form of skill learning.

Vatterott, D.B., & Vecera, S.P. (2012). Experience-dependent attentional tuning of distractor rejection. *Psychonomic Bulletin & Review, 19*(5), 871-878. doi: 10.3758/s13423-012-0280-4

Vecera, S.P., Cosman, J.D., Vatterott, D.B., & Roper, Z.J.J. (2014). The control of visual attention: Toward a unified account. In Brian H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 60, pp. 303-347). Academic Press. doi: 10.1016/B978-0-12-800090-8.00008-1

Vickery, T.J., Sussman, R.S., & Jiang, Y.V. (2010). Spatial context learning survives interference from working memory load. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1358-1371. doi: 10.1037/a002558

Vogel, E.K., & Machizawa, M.G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428*(6984), 748-751. doi: 10.1038/nature02447

Vogel, E.K., Woodman, G.F., & Luck, S.J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance,32*(6), 1436-1451. doi: 10.1037/0096-1523.32.6.1436

Walthew, C., & Gilchrist, I.D. (2006). Target location probability effects in visual search: An effect of sequential dependencies. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1294-1301. doi: 10.1037/0096-1523.32.5.1294

Wilder, M.H., Mozer, M.C., & Wickens, C.D. (2011). An integrative, experience-based theory of attentional control. *Journal of Vision, 11*(2), 8. doi: 10.1167/11.2.8

Wolfe, J. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202-238. doi: 10.3758/BF03200774

Wolfe, J. M., Horowitz, T.S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research, 44*(12), 1411-1426. doi: 10.1016/j.visres.2003.11.024

Woodman, G.F., & Luck, S.J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review, 11*(2), 269-274. doi: 10.3758/BF03196569

Woodman, G.F., & Luck, S.J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search. *Journal of Experimental Psychology: Human Perception and Performance, 32*(2), 363-377. doi: 10.1037/0096-1523.33.2.363

Woodman, G.F., Luck, S.J., & Schall, J.D. (2007). The role of working memory representations in the control of attention. *Cerebral Cortex*, *17*(suppl 1), i118-i124. doi: 10.1093/cercor/bhm065

Woodman, G.F., Vogel, E.K., & Luck, S.J. (2001). Visual search remains efficient when visual working memory is full. *Psychological Science, 12*(3), 219-224. doi: 10.1111/1467-9280.00339

Zehetleitner, M., Goschy, H., Müller, H.J. (2012). Top-down selection of attention: It's gradual, practice-dependent, and hierarchically organized. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 941-957. doi: 10.1037/a0027629