

Effects of Expectation, Experience, and Environment on Visual Search

by

Mathias S. Fleck

Department of Psychology & Neuroscience
Duke University

Date: _____

Approved:

Stephen Mitroff, Supervisor

Amy Needham

David Madden

Elizabeth Marsh

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctorate
of Philosophy in the Department of
Psychology & Neuroscience in the Graduate School
of Duke University

2009

ABSTRACT

Effects of Expectation, Experience, and Environment on Visual Search

by

Mathias S. Fleck

Department of Psychology & Neuroscience
Duke University

Date: _____

Approved:

Stephen Mitroff, Supervisor

Amy Needham

David Madden

Elizabeth Marsh

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctorate in the Department of
Psychology & Neuroscience in the Graduate School
of Duke University

2009

Copyright by
Mathias S. Fleck
2009

Abstract

A pervasive aspect of daily life is searching for a specific target amongst an array of distracting items. Studying such visual searches offers a useful and powerful tool for revealing the underlying aspects of visual attention. Understanding how factors influence accurate target detection serves to both enhance real-world search tasks and inform basic cognitive psychology. The goal of the research presented herein is to examine the effects of expectation, experience, and environment on search behavior. The experiments are conducted in controlled laboratory environments, but are designed to simulate real-world searches, with the express goal of informing the implementation of search tasks in everyday life. First, *expectation* is explored by manipulating target prevalence and measuring the resultant change in behavior as participants' biases shift. Second, *experience* is tested by comparing individuals with and without extensive video game exposure, specifically on their susceptibility to the pressures of rare target search. Lastly, *environment* is examined by utilizing multiple simultaneous targets. This manipulation has been shown to induce errors in radiology, and here the generality of this effect is explored to establish the various pressures to which it is sensitive. Collectively, these data serve to inform how different influences modulate visual search performance, and the results can directly inform the training, recruitment, and execution of real-world search tasks such as those in radiology, cytology, and airport security.

Dedication

For my parents, Carl and Rosa, and in memory of my grandma, Gaynelle Seals (1924-2008).

Contents

Abstract	iv
List of Tables	x
List of Figures	xi
Acknowledgements	xii
Chapter 1. Introduction.....	1
1.1 What is visual search?.....	2
1.1.1 Visual search as an operational measurement of visual attention	3
1.1.2 Applied search in professional and medicinal applications	4
1.1.3 Applied search in military and security sectors.....	6
1.2 Review of visual search in cognitive psychology	7
1.2.1 Measuring search in the laboratory	7
1.2.2 Feature integration theory.....	8
1.2.3 Parallel and serial processing	9
1.2.4 A new model: guided search	10
1.2.3 Guiding attributes	12
1.2.4 Bottom-up and top-down guidance	13
1.2.5 Memory processes and search history	14
1.3 Current topics in visual search	16
1.3.1 Target prevalence and expectation	16
1.3.1.1 Vigilance studies	17

1.3.1.2 Prevalence studies in radiology	18
1.3.1.3 Prevalence effects in cognitive studies.....	19
1.3.2 Video game experience.....	20
1.3.2.1 Video game expertise influences attention and perception.....	21
1.3.2.2 Video games and real-world tasks	22
1.3.2.3 Possible mechanisms of video game effects.....	22
1.3.3 Extra targets in the search environment	23
1.3.3.1 Radiology and the satisfaction of search effect.....	24
1.3.3.2 Generality of SOS effect?.....	24
Chapter 2. Expectation and rare target search.....	26
2.1 Methods	27
2.1.1 Participants and stimuli.....	27
2.1.2 Procedure.....	29
2.2 Results and Discussion	30
2.3 Additional findings in rare target search.....	37
2.3.1 Speed/accuracy tradeoff or criterion shift?	37
2.3.2 Change in quitting threshold.....	40
2.3.3 Correction does not always eliminate the prevalence effect.....	42
Chapter 3. Video game experience improves rare target search.....	45
3.1 Methods	48
3.1.1 Participants and stimuli.....	48
3.1.2 Procedure.....	50

3.2 Results	51
3.2.1 Visual search accuracy	51
3.2.2 Response times	53
3.2.3 Speed-accuracy tradeoffs.....	54
3.3 Discussion.....	56
3.3.1 Video gamers excel at rare target visual search	56
3.3.2 Role of VGP motivation in enhanced rare search performance.....	58
3.3.3 Implications for prior and future video game research.....	59
3.3.4 Conclusions and broader implications.....	60
Chapter 4. Extra targets in the visual search environment	62
4.1 Experiments 1-3: Examining the effect of salience and expectation on SOS.....	65
4.1.1 Methods	66
4.1.1.1 Participants and stimuli	66
4.1.1.2 Procedure	68
4.1.1.3 Planned analyses.....	69
4.1.2 Results and Discussion	71
4.2 Experiments 4-5: Examining effects of decision-making processes on SOS	74
4.2.1 Methods	75
4.2.1.1 Participants and stimuli	75
4.2.1.2 Procedure	76
4.2.2 Results and Discussion	78
4.3 Experiments 6-8: Examining the effect of time pressure on SOS.....	80

4.3.1 Methods	81
4.3.2 Results and Discussion	82
4.4 Experiments 9-10: Examining how identical salience in additional targets modulates SOS	83
4.4.1 Methods	85
4.4.2 Results and Discussion	85
4.5 Experiment 11: Examining the effect of instruction on SOS.....	87
4.5.1 Methods	88
4.5.2 Results and Discussion	88
4.6 Experiments 12-13: Examining the interaction of time and reward pressures on SOS.....	89
4.6.1 Methods	91
4.6.2 Results and Discussion	93
4.7 General Discussion.....	94
4.7.1 Summary	95
4.7.2 Implications for SOS in radiology	97
4.7.3 Implications for cognitive psychology	99
4.7.4 Implications for luggage screening and other real-world searches	101
4.7.5 Conclusion.....	103
4.8 Appendix: Response time data.....	104
Chapter 5. Conclusion	106
References	114
Biography	126

List of Tables

Table 1: Experimental parameters and accuracy data for SOS Experiments 1-13.	73
Table 2: Response time data for SOS Experiments 1-13.	105

List of Figures

Figure 1: Find the blue circle. (a) Example of an efficient search. (b) A less efficient search.	10
Figure 2: Completion of occluded figures can occur preattentively. In (a), the partly occluded circle can be searched for efficiently, unlike (b), which perceptually has an identical pac-man shape but lacks occlusion cues.	11
Figure 3: Sample search array. Observers searched for a tool amidst randomly rotated items from other categories.	28
Figure 4: Miss rates by target prevalence. *Low-prevalence signifies 1% for Wolfe et al. and 2% for the current experiment.	31
Figure 5: Low prevalence response time data from the No-Correction condition organized by ordinal relationship to a target-present trial.	33
Figure 6: Post-correction miss rates for low-prevalence trials for VGPs and NVGPs at each Prevalence level.	53
Figure 7: Response times for low-prevalence trials for VGPs and NVGPs at each Prevalence level.	54
Figure 8: Speed and accuracy are plotted for each participant at (a.) High prevalence, (b.) Medium prevalence, and (c.) Low prevalence.	55
Figure 9: Sample search display for SOS Experiments 1-3, 6-11.	68
Figure 10: Sample search display for SOS Experiments 4 and 5.	76
Figure 11: Sample search display for SOS Experiments 12 and 13.	92

Acknowledgements

Along the way, there were many false starts, mistakes, insecurities, null findings, freak outs, and one total reboot. I am thankful to many folks. First and foremost, this work would be nonexistent if not for the persistent guidance of my advisor, Steve Mitroff. Steve is the ideal mentor: He gave me a vast freedom to explore the topics I found most interesting, then totally engaged those projects and supported me through to completion. He offers the finest example of how to think and write about science, and how to present it with total lucidity. I am honored to be his first graduate student.

I thank Roberto Cabeza, whose lab I initially joined, for being incredibly supportive as I tried to find the right path. I am also grateful to all of the professors at Duke who served on my committee at one time or another, every one of whom set aside time to help me with my problems, both scientific and otherwise: Amy Needham, Beth Marsh, Ian Dobbins, David Madden, and Kevin Pelphey. I am especially grateful to Amy for being there for me during an extremely shaky first year. Her caliber of genuine understanding is rare in academia. I also adore the warmth, kindness, and tirelessness of Fonda Anthony, who always checked in to make sure I was OK. I thank the members and friends of our lab for helpful conversation and assistance: Melissa Bulkin, Laura Sestokas, Sarah Donohue, Joe Harris, Ricky Green, Viktoria Elkis, Brittany James, Jordan Axt, Greg Appelbaum, Yoni Mazuz, and George Alvarez.

I chose Duke over other schools because of the awesome people I met during recruitment. These people and others who have come through the department are now some of my closest friends; they truly made it a home for me: Jen Talarico, Simon Tonev, Chrissy Camblin, Michele and Wilson Diaz, Dan Dillon, Chris MacDonald, Heather Rice, Mike Studson, Umay Suanda, Evan Maclean, Jason Arita, Kait Clark, Betsy Holmberg, and Katrina Poetzl. As many of these moved on from Duke, their absence hammered home their significance in making my graduate life a pleasure.

I thank Jessica Cantlon, who is single-handedly responsible for keeping me in grad school. Her brilliance and friendship challenged me at a deeper level than any class ever would. I also thank Kris Riddle, who is the funniest and sweetest roommate and friend anyone could ask for. For championing me like no one else, I am particularly grateful to Caroline Cozza. Her compassionate, generous, and unflinching support of me during the writing of this dissertation was simply invaluable.

I thank too the friends whose humor, smarts, and sheer ownage factor shaped my earlier years and prepared me for this: Pat Burns, Marty Golando, Mike Carpenter, Jason Burge, Ryan Phillips, Brian Showers, Phil Weintraub, Will Wilson, Britt Crawford, Rob Hoff, and Jorge Araujo. Their ongoing role in my life is irreplaceable.

Finally, I am grateful to my wonderful family. I thank my brother and sisters and all their kids for the vital, recharging reprieve they offered for a few days each year at the holidays: the Flecks (Maggie, Joe, Elsa, Monti, and Dahlia) and the Haddons (Mary,

Steve, Josh, Ben, and Eliza). I cherish those visits and look forward to many more. Above all, I thank my parents, Rosa and Carl, for their love, and for their endless desire to make everything absolutely perfect for me.

Chapter 1. Introduction

Visual attention is a fundamental cog in the machinery of human cognition.

Faced with a barrage of incoming information, we rely on visual attention to give order to our environment. After parsing a scene into its distinct components, our visual system directs attention to specific elements for additional processing. This act of selecting particular items, features, or locations in space can be referred to as *visual search*, and it is a task we perform perpetually. From reading a book to rummaging through the refrigerator, visual searches are continually conducted, with a wide range of necessary effort and specificity of targets. It is the pervasiveness of this basic cognitive skill that motivates the present research on the factors that modulate visual search behavior.

The goal of the research reported here is to explore the impact of *expectation*, *experience*, and *environment* on visual search in healthy human subjects, with the express goal of improving performance in critical real-world search tasks. Chapter 1 begins by illustrating some of the real-world examples that motivate this work and offers further discussion about why it is important to study visual search. Many of the fundamental properties critical to visual search have already been richly explored in cognitive psychology (see Wolfe, 2007), and Chapter 1 reviews this groundwork before turning to a discussion of the current topics in visual search that are experimentally tested in the remainder of the dissertation.

Chapter 2 explores how the prevalence of sought-after targets influences *expectation* about the search task and subsequently changes detection performance. Chapter 3 extends the prevalence work by examining the interaction between expectation and *experience*. This experiment specifically examines how the experience of video game play, known to have generalized effects on perception and attention, modifies speed and accuracy in rare target search. Chapter 4 focuses on effects of the search *environment*, in this case how the presence of additional targets in the search space can modulate accuracy. Beyond establishing the general effects of multiple target search, this chapter presents thirteen experiments which illuminate a host of pressures to which multiple-target search is sensitive. Chapter 5 concludes with a brief discussion of how these external influences add to the visual search literature and carry significant implications for the design and implementation of real-world search tasks.

1.1 What is visual search?

When we walk into a grocery store with a list of items to purchase, we are faced with the daunting prospect of sorting through an overwhelming array of products to find the desired subset of those products on our list. As we scan a wall of brightly-labeled products, all carefully designed to maximize the capture of attention, we engage a prototypical visual search process to match our mental templates of known color and shape patterns to this streaming input of visual information. Upon fixating and

recognizing the bright blue color swathes and anthropomorphic tiger that typify a box of Frosted Flakes, we grab the item and move down our list.

As mundane a task as this is, the successful detection of the target and the time required to seek it out depends directly on a vast array of factors beyond the simple low-level pattern matching involved in color and shape discrimination. A brief consideration of these influences illustrates the multitude of external variables which interact to modulate visual search behavior, including: the similarity and proximity of other brands of cereal, the number of Frosted Flakes boxes available on the shelf relative to the number of non-Frosted Flakes boxes, the amount of background noise, the familiarity of the shopper with the particular store, the knowledge about whether the store carries the particular brand, and whether or not the shopper is simultaneously looking for other cereals or products. These same elements modulate the detection performance in a host of other everyday search tasks, such as when we pick out a familiar face in a crowded bar, locate our car in a packed shopping mall parking lot, or try to figure out where we are on a map.

1.1.1 Visual search as an operational measurement of visual attention

Broadly speaking, the cognitive process of visual search concerns how we deploy visual attention to organize perception. In the words of William James (1890), “My experience is what I agree to attend to. Only those items which I notice shape my mind – without selective interest, experience is an utter chaos.” More specifically, visual search

can be operationally defined as the allocation of visual attention to specific features or locations in space which are selected for additional processing (i.e., “signal”) while deprioritizing, limiting, or altogether excluding the analysis of other input (i.e., “noise”) that is less relevant to the search at hand. A typical search likely consists of several consecutive deployments of attention until either the target has been found or once the searcher abandons the search.

Depending on the properties of the target signal and the surrounding noise in which it is embedded, the process can sometimes be carried out instantaneously and automatically (i.e., a “parallel” search), or it can require an item-by-item, more effortful scan (i.e., a “serial” search) to locate the target. This search process is driven in part by the top-down boosting of stimulus properties relevant to the search (e.g., Melcher, Papathomas, & Vidnyánszky, 2005), a process referred to as “guidance” of attentional selection (Wolfe, 2007), although there is also evidence to suggest that the search process involves the inhibition of irrelevant stimulus attributes (e.g., Most et al., 2001). Other basic properties of visual search have been explored experimentally and are reviewed in Section 1.3, but it is worth first reflecting on why the topic deserves scrutiny by considering the broader social contexts in which visual searches occur.

1.1.2 Applied search in professional and medicinal applications

In the professional domain, many jobs critically depend on the accurate and swift execution of a visual search. Charting the parameters to which visual search is sensitive

and working to optimize these factors thus carries potentially significant social importance. The lifeguard in charge of a summer pool must be alert for the flailing of swimmer in trouble (e.g., Lanagan-Leitzel & Moore, 2008), a rare occurrence potentially complicated by the visual similarity to the splashing of other pool-goers. Termite inspectors are trained to identify subtle indications in the building structures and soil patterns which evidence termite infestation, and although such instances are infrequent, the harms associated with missing the subtle clues are professionally disastrous. The rarity of pool emergencies and termite infestations highlights a characteristic common to many real-world searches: *low incidence rates for target events*. The effect of target prevalence is directly addressed in Chapter 2 and is a predominant theme throughout this dissertation.

Industrial inspection is another professional sphere reliant on the efficiency of visual search, demanding that inspectors notice small deviations from normalcy that necessitate correction or removal of an item from production. Indeed, this motivation has fostered an entire branch of research within industrial psychology to increase inspection efficiency (Münsterberg, 1913; Rose, 1975; Wyatt & Langdon, 1932; Wyatt, Langon, & Stock, 1938), much of which has been concerned with the influences of boredom and monotony on inspection performance (for review, see Davies, Shackleton, & Parasuraman, 1983). This focus not only emphasizes the repetitiveness of such tasks, but as well the low frequency with which aberrations occur.

Pressures associated with target rarity are potentially compounded even further by the complexity of the search task, and nowhere is this complexity likely greater than in the domain of medical imaging, as in radiology and cytology. The global difficulty of these search tasks, the possibility of multiple targets to be found, the limitations of imaging technology, and the dire consequences of any missed targets together emphasize the demand for extensive training and a developed intuition to properly detect in radiographic images an unexpected abnormality in a routine exam, or to efficiently scan a Papanicolau smear for the presence of rare abnormal cells. Accordingly, numerous studies have been conducted to understand errors in medical image searches, some of which are reviewed in Section 1.3.3. Such work directly motivated the studies presented in Chapter 4 that explore the generality of multiple target effects noted in medical studies.

1.1.3 Applied search in military and security sectors

Charting the parameters of visual search also carries significant implications for military search tasks. Experimental psychologists were recruited early in World War II to study information processing akin to visual search, such as how to optimize the detection of weak signals by radar operators (Fitts, 1947). Although much of this early work was hampered by overly theoretical approaches, psychologists adapted to designing studies for application in real-world implementations, birthing a new branch

of study known as engineering psychology (Chapanis, 1999), and it is this practical ideology which motivates the studies presented in Chapters 2-4.

Lastly, by exploring visual search here, we can directly inform the kinds of search tasks typical of security screenings, such as those in airport luggage X-raying. Airport screeners are trained to consider multiple factors which might hinder successful visual search, including item superpositioning (overlapping), atypical viewpoint rotation, and overall clutter/image complexity (Schwaninger, 2005). Yet, a multitude of other pressures may also be at work, including biased expectancy (a gun is an extremely unlikely event), insufficient training (not every kind of gun was observed in training), psychosocial stressors (a long line might increase pressure to search quickly), and the breadth of the target pool (adding liquids and gels to the list of dangerous items potentially reduces detection of guns and bombs). Work examining some of these factors is reported herein.

1.2 Review of visual search in cognitive psychology

1.2.1 Measuring search in the laboratory

The basic properties of visual search have been studied extensively and modeled in controlled laboratory experiments. A typical laboratory study consists of a series of consecutive trials performed on a computer in which participants attempt to locate a pre-specified target amidst an array of distractors and press a button to indicate either the presence or absence of the target. In most studies, half of the trials contain a target,

and half of the trials do not. The time taken to find the target (i.e., response time) is recorded for each trial, and since accuracy on most tasks is quite high, studies frequently examine the impact of experimental manipulations on response time as the primary measurement. Increasing the number of distractors (i.e., set size) usually increases the search time, and it typically takes longer to determine that a target is absent than to identify its presence (Chun & Wolfe, 1996). A useful standard for comparing visual search properties is to plot the response times at several set sizes and thus determine for a given task the “search slope,” or the linear rate of search slowdown which can be described in terms of seconds per the number of items present in the display.

1.2.2 Feature integration theory

Current models of visual search owe much to Treisman’s feature integration theory (FIT; Treisman & Gelade, 1980; Treisman, 1982, 1988), which was the standard for search models for the past two decades. FIT builds upon the proposal by Neisser (1967) that visual processing should be divided into “preattentive” and “limited capacity” processes. The FIT model is a linear, two-stage process wherein basic features (e.g., color or size) are first processed and parsed in a “preattentive” fashion, without effort, automatically, and efficiently. This initial process of feature generation is followed by a slower process that requires directed attention to “bind” features into objects (Kahneman, Treisman, & Gibbs, 1992) and select some subset of those objects for further,

more advanced processing. This attentive binding is considered the primary bottleneck in the visual search process.

FIT is supported in part by studies demonstrating errors of “illusory conjunctions” (Treisman & Schmidt, 1982), in which the rapid presentation of stimuli consisting of two or more basic features will lead observers to report mismatches between these basic features. For example, presenting a red square and a blue circle may be subsequently reported as a blue square and a red circle. The idea here was that the rapid presentation only allows preattentive analysis to gather the basic features of the display (red, blue, circle, square), but did not allow for enough time for directed attention to bind those features into the objects appropriately (e.g., a red square). Treisman and Schmidt (1982) also suggested that the occurrence of illusory conjunctions may itself be a criterion to determine what constitutes a basic feature.

1.2.3 Parallel and serial processing

In a simplified view of visual processing, the two-stage FIT model predicts two broad categories of searches: parallel and serial. If a target is distinct from distractors along one or more basic featural dimensions (i.e., finding a blue circle amidst red squares; Figure 1a), the target can typically be located instantly and automatically (i.e., a “pop-out” effect), with little effect on response time of increasing set size, and such a search is described as a parallel search. Alternately, if a target is specified by the conjunction of multiple features and if some basic features are shared with distractors

(i.e., finding a blue circle amidst blue and red squares; Figure 1b), the search requires an item-by-item, serial analysis of each distractor to determine whether it is a target. The time to conduct a serial search is therefore dependent on set size and is a much slower search. Although this strictly dichotomous view of visual search is no longer considered to be an accurate characterization (see Townsend, 1990; Wolfe, 1998), the parallel/serial division continues to offer two endpoints of a spectrum along which we can plot, for any given search, the efficiency to find the target as set sizes varies.

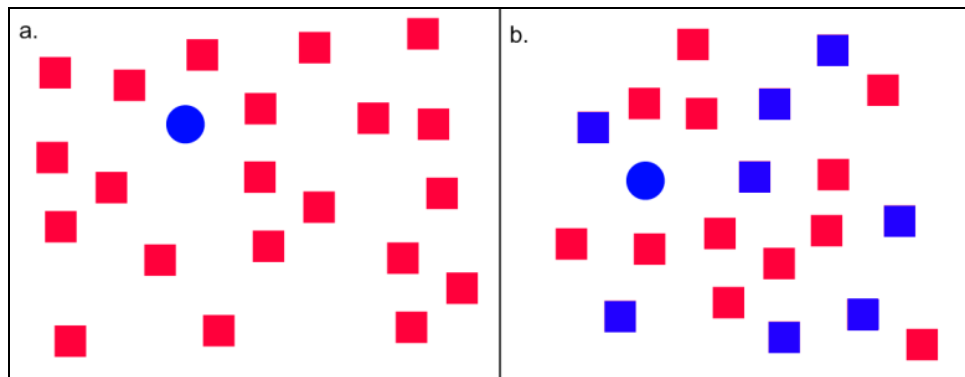


Figure 1: Find the blue circle. (a) Example of an efficient search. (b) A less efficient search.

1.2.4 A new model: guided search

While FIT neatly captures many of the characteristics of visual search, the linear two-stage model of FIT suffers from accumulating evidence that much more is happening preattentively than FIT allows (Wolfe, 1998; Wolfe & Horowitz, 2004). For example, searches for partly occluded objects (Figure 2a) were shown to be more efficient than identical shapes without occlusion cues (Figure 2b), suggesting that the visual system automatically completed the partly hidden shapes very early in

processing (Rauschenberger & Yantis, 2001; Rensink & Enns, 1998). However, FIT supposes that directed attention is a necessary step to bind the visual features indicating overlap into distinct objects. Additionally, evidence for object-based attention, wherein an entire object is selected for an attentional boost rather than just a spatial location (Duncan, 1984; Egly, Driver, & Rafal, 1994), suggests that some preattentive object binding must occur.

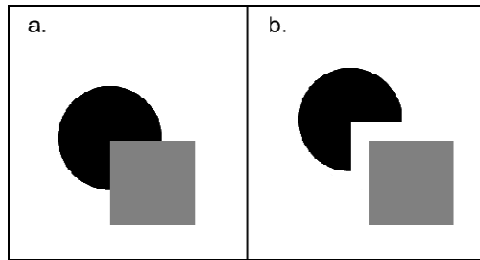


Figure 2: Completion of occluded figures can occur preattentively. In (a), the partly occluded circle can be searched for efficiently, unlike (b), which perceptually has an identical pac-man shape but lacks occlusion cues.

A more recent update on FIT is a model proposed by Jeremy Wolfe called “guided search” (2007). It too adopts a two-stage preattentive and attentive division like FIT, but guided search de-emphasizes the strictly linear structure of basic-features-into-objects by instead treating basic features as guiding attributes which can direct the deployment of attention, but function more as a control module than a necessary precursor to objecthood. This model allows for both bottom-up and top-down input by positing a coarse, preattentive representation extracted from bottom-up stimulation upon which top-down mechanisms can bias a particular feature or sets of features for future selection as an object. Importantly, the model does not necessitate that objecthood

can only arise in the latter stage; guided search allows that even prior to focused attention, preattentive structures can be quite complex (i.e., “proto-objects,” Rensink, 2000), although without attention the representations are temporally volatile.

1.2.3 Guiding attributes

Given the special role of guiding attributes, previous work has attempted to map out these specific properties that are available to preattentive access (Wolfe, 1998, 2007; Wolfe & Horowitz, 2004). Guiding attributes are primarily identified by their ability to generate efficient (i.e., parallel) searches. This is usually experimentally determined by manipulating set size and finding little or no effect on response times (i.e., yielding a flat or shallow search slope). However, this is not a sufficient rule for inclusion as a guiding attribute (Wolfe & Horowitz, 2004). Another important characteristic used to identify a guiding attribute is whether it demonstrates “search asymmetry” (Treisman & Gormican, 1988; Wolfe, 2001), which refers to the tendency for a feature to be more notable in its presence than in its absence. For example, it is easier to find a moving target amidst stationary distractors than vice versa. As described in the previous section if a feature exhibits illusory conjunctions, this may also provide converging evidence that it represents a basic feature/guiding attribute.

A bulk of evidence reveals four guiding attributes that handily pass these criteria: color (Bundesen & Pedersen, 1983; Duncan, 1989), motion (Dick, Ullman, & Sagi, 1987), orientation (Foster & Ward, 1991), and size (Quinlan & Humphreys, 1987;

Treisman & Gelade, 1980). A number of other features offer some evidence for preattentive processing, but are not as clearly fundamental in guiding access to the binding stage. Wolfe and Horowitz (2004) offer a comprehensive list of these possible guiding attributes as well as an assessment of the strength of the evidence to support each of the features. These possible guiding attributes include luminance, shape, curvature, novelty, familiarity, and closure, among others.

1.2.4 Bottom-up and top-down guidance

In parallel searches, the guiding attributes of particular targets are so distinct from the surrounding distractors that attention may be automatically captured. In other cases, items in a scene may be too similar to evoke automatic attention capture. This distinction of a target from its local background and its heterogeneity with respect to nearby distractors represents the *saliency* of the target. An item may have increased detectability purely as a function of unique disparity from distractors along one or more guiding attributes, and this reflects bottom-up saliency (Duncan & Humphreys, 1989). However, the goals that the viewer brings to the search also modulate the saliency of items (Yantis, 1998). For example, if one is looking for a key on a cluttered desk, the characteristics of “shiny” and “silver” are made salient by top-down effortful search.

Naturally, there will arise situations in which bottom-up and top-down factors conflict. In the previous example, a bright highlighter on the desk may draw attention automatically, in spite of not matching the predetermined goals of the search.

Depending on the parameters of the search, top-down control can sometimes override salient distractors (e.g., Leber & Egeth, 2006), yet at other times bottom-up salience may capture attention in spite of the searcher's high-level goals (Remington, Johnston, & Yantis, 1992).

1.2.5 Memory processes and search history

Beyond the memory processes required for recognizing targets, there are other roles for memory in determining search behavior. For tasks involving repeated and consecutive searches, the recent detection of a target on one trial has been shown to influence performance on subsequent trials. When a target defined by a particular basic feature repeats within the next few trials, a priming mechanism automatically enhances the "pop-out" effect for that feature (Maljkovic & Nakayama, 1994), and this priming benefit has been demonstrated for searches guided by top-down factors as well (Hillstrom, 2000). Interestingly, in special cases of search, memory for recent targets can actually have a negative top-down effect on overall accuracy. Across a series of trials, if target prevalence is very low (e.g., 1 out of every 100 trials contains a target), then the successful detection of a target leads searchers to temporarily speed up on subsequent trials in what appears to be a kind of "gambler's fallacy," and this faster responding predictably leads to missed targets (Fleck & Mitroff, 2007; Wolfe, Horowitz, & Kenner, 2005; see Chapter 2 for more details).

Another significant effect of memory across a search session can be observed if the spatial configuration of targets and/or distractors repeats on future trials. Even when searchers are not explicitly aware of regularities in the spatial pattern of targets and distractors, visual attention can be guided to predictive locations in a process known as contextual cueing (Chun & Jiang, 1998). Importantly, the ability to implicitly learn spatial relationships suggests that powerful, adaptive heuristics may be at work “behind-the-scenes” of visual search to optimize search behavior, and it is critical to establish what other heuristics may be operating automatically. Chapter 4 explores the possibility of another automatic search heuristic whereby the detection of one target may trigger early termination of a search, preventing an exhaustive search.

Lastly, memory may play a within-trial role in determining the serial scan path of a particular trial as individual locations are explored, but this point is not well established. Horowitz and Wolfe (1998) demonstrated that randomly rearranging items within a trial generated no change in efficiency, suggesting that the search mechanism does not rely on short term memory to encode previously-attended locations. However, others have argued that some searched locations are indeed inhibited for future search (Kristjansson, 2000), similar to the mechanism of inhibition of return (Posner & Cohen, 1984), wherein recently attended locations or objects are deprioritized for future search. One compromise between the two positions (“visual search has no memory” versus “inhibitory tagging of all previously attended locations”) poses a hybrid model related

to optimal foraging behavior (Klein & MacInnes, 1999), wherein some subset of previously attended locations are tagged for inhibition, but that this process is not necessary or complete.

1.3 Current topics in visual search

As the previous section reviewed, cognitive psychology has established many of the fundamental properties of visual attention and visual search, delineating which features guide search, and how (e.g., Wolfe, 2007). Yet, while some previous work examined external influences like search history (e.g., Maljkovic & Nakayama, 1994), the bulk of work has tested within-display manipulations, and several “higher order” factors remain underexplored. How is visual search affected by factors beyond the specific visual information present within a given scene? A wide host of external pressures may interact to modify performance of visual search, and this introductory chapter concludes by focusing specifically on three of these influences: expectation, experience, and the search environment.

1.3.1 Target prevalence and expectation

Laboratory studies of visual search typically include targets on half of all trials, and participants quickly recognize that any given trial has an equally likely chance of having a target or not. However, real-world searches are rarely so balanced. In airport security, for example, luggage screeners view thousands of X-rays in a week, yet the sighting of a dangerous item such as a gun or knife may happen only once a year or

even less frequently (Rubenstein, 2001). Thus, one current issue in visual search is exploring how manipulating *target prevalence* modulates searcher expectation and subsequent detection performance. Specifically, how do response times and detection accuracy change as a function of target rarity? This issue has been raised in multiple domains, including vigilance studies, radiology work, and cognitive psychology.

1.3.1.1 Vigilance studies

Although target prevalence has only recently been scrutinized in traditional visual search paradigms (e.g., Fleck & Mitroff, 2007; Wolfe et al., 2005; Wolfe et al., 2007), the topic bears much resemblance to early studies of vigilance borne out of war efforts to improve military performance in World War II. Psychologists were recruited to quantify and reduce the error rates of monotonous tasks such as radar operators, and early studies indicated that even within the first hour of a monitoring task, detection performance would drop around 10% and another 5% after another hour (Mackworth, 1950). Parasuraman and Davies (1976) examined the interaction between elapsed time-on-task and probability of a target event, and demonstrated both a main effect on accuracy for probability (lower probability decreased performance) and a main effect of time elapsed (accuracy decreasing over time). For a review of these early studies of vigilance, monitoring, and boredom, see Davies, Shackleton, and Parasuraman (1983).

How do vigilance tasks map onto modern visual search paradigms? Vigilance studies typically consisted of a monitoring task in which events occur at unknown

intervals, in contrast with a visual search study in which each trial demands a separate response of absent or present. Although boredom and monotony are potential factors in any long study of visual search, this critical difference in responding may demand separate consideration of these paradigms. Passively staring at a radar screen, waiting for a blip to appear may be quite different from actively scanning and individually responding to consecutive presentations of cluttered arrays. The differing response mechanisms may be responsible for one noticeable and likely significant difference between the paradigms: whereas response times typically increase over time in a vigilance/monitoring task (Buck, 1966), responses speed up over the course of a rare target paradigm (e.g., Wolfe et al., 2005), suggesting a potentially different cognitive mechanism at work.

1.3.1.2 Prevalence studies in radiology

In routine radiological examinations, the incidence rate for abnormalities is extremely low. The actual rate varies considerably depending on the demographics of the population, the procedure being used, the body part being examined, and the nature of the screening, but in all cases a missed identification carries potentially severe consequences. Furthermore, the rate of missing targets has been consistently shown to be alarmingly high, typically around 30% (Berlin, 1994; Kundel, 1989; Renfrew, Franken, Berbaum, Weigelt, & Abu-Yousef, 1992). As a consequence, work within radiology has

attempted to address whether low target prevalence may be responsible for high rates of error.

In one study of pulmonary arteriograms, the six observers yielded significantly more accurate diagnoses for pulmonary emboli when the prevalence in the samples was 60% versus when it was only 20% (Eggin & Feinstein, 1996). More recently, however, in a comprehensive radiological study of low prevalence in the diagnosis of posteroanterior chest images, Gur and colleagues (2003) concluded that a prevalence effect was not driving miss rates in their controlled laboratory study, which varied prevalence from 28% to 2%. Importantly, this study utilized more observers and lower incidence rates than Eggin and Feinstein (1996), which may better map onto actual rates of screenings or routine examinations. An additional study by the same group (Gur et al., 2007) demonstrated an influential effect of prevalence expectations on confidence ratings following target identification, in which decreasing prevalence tended to increase confidence ratings, yet again the data indicated no detrimental effect on accuracy.

1.3.1.3 Prevalence effects in cognitive studies

In view of the conflicting reports from radiology and the possibility that highly trained radiologists may conduct searches differently than a “typical” searcher, who may be more or less prone to a prevalence effect, this issue has recently been raised in traditional visual search studies within cognitive psychology to look for a potentially

generalized influence of target prevalence on search behavior. A recent laboratory study (Wolfe et al., 2005) revealed a striking effect of decreasing prevalence in typical visual search studies: observers missed only 7% of targets when prevalence was high (targets present on 50% of trials), but missed an alarming 30% when prevalence was low (1% prevalence). The authors associated the increasing rate of errors to a speeding of responses, and the finding carries potentially significant implications for comparable search tasks, such as airport security screening.

However, it is important to fully establish the source of the drastic increase in errors. Importantly, what happens to target information on the miss trials in which participants responded too quickly? Do high miss rates in rare target search represents errors of perception or errors of action? Research reported in Chapter 2 follows up on the Wolfe et al. (2005) study to better understand the origin of the low-prevalence effect, with the goal of reducing similar errors in real-world tasks such as airport security screening.

1.3.2 Video game experience

Another topic addressed by the present work is the role of *individual differences* in determining search behavior. How do different experiences or traits influence visual attention? Previous studies have examined how different populations differ across visual search tasks for two reasons: to inform the basic principles of visual attention, and to characterize the cognitive differences between groups. For example, it has been

demonstrated that autistic individuals demonstrate significantly better visual search performance than non-autistic controls, offering insight into the cognitive peculiarities of autism (O’Riordan, Plaisted, Driver, & Baron-Cohen, 2001; Plaisted, O’Riordan, & Baron-Cohen, 1998). Visual search has also been used as a tool to compare young and elderly subjects and to demonstrate that age-related decline in selective attention is not driven by attentional capacity limitations (Madden & Langley, 2003).

There has been little work examining video game expertise and visual search (but see Castel, Pratt, & Drummond, 2005), yet a growing body of research illustrates the potentially significant effects that video gaming may have on other tasks of attention and perception. Furthermore, there is now evidence to suggest that video game playing may directly impact performance on many real-world tasks as well.

1.3.2.1 Video game expertise influences attention and perception

Previous research has shown that expert videogame players (VGPs) differ from non-videogame-players (NVGPs) on a variety of attention and perception tasks. For example, VGPs demonstrate enhanced visual acuity (Green & Bavelier, 2007), maintain a wider field of attention (Green & Bavelier, 2003; 2006a), have a higher resolution of temporal attention (Green & Bavelier, 2003), can better track multiple moving objects (Green & Bavelier, 2006b), and they generally respond faster (e.g., Castel et al., 2005; Green & Bavelier, 2007). Importantly, these benefits represent skills enhanced via videogame experience—when NVGPs are trained in videogame play, they demonstrate

VGP-like advantages (e.g., De Lisi & Cammarano, 1996; Gopher, Weil, & Bareket, 1994; Green & Bavelier, 2003, 2006a, 2007).

1.3.2.2 Video games and real-world tasks

The growing body of empirical evidence for VGP benefits has recently extended to more complex tasks with real-world attentional demands. Laparoscopic surgeons with video game experience outperform those without video game experience (Rosser et al., 2007). In the study, surgeons who played video games at least three hours per week in their past were 27% faster than non-gamer colleagues, with 37% fewer errors.

Video games have also been used as a training tool to improve performance in military training. Games such as *America's Army* have been specifically designed to teach proper procedures and to improve response times in critical situations.

1.3.2.3 Possible mechanisms of video game effects

Video game experience has been shown to influence attention and perception and to modify performance of real-world, attention-demanding tasks, yet there has been little discussion of the specific mechanisms driving the benefits. Importantly, although video game experience has been associated with perceptual learning (e.g., Green & Bavelier, 2003), one critical difference is that perceptual learning effects are extremely specific to the learned task (e.g., Fiorentini & Berardi, 1980), in contrast to the task-generalizability of video game playing effects.

It is possible that video game effects reflect a new form of generalized perceptual learning. Such a mechanism implies that the experience causes low-level modulation of visual processing in which video game playing creates enhanced attentional capacities and spatio-temporal processing abilities. An alternate hypothesis is that video game experience drives higher-level differences, such as video gamers possessing higher motivation to perform a task well or applying better strategies to yield enhanced performance over non-gamer participants.

These mechanism possibilities have not previously been explored in video game experience research. Chapter 3 presents research that addresses this mechanistic question by examining how video game experience influences rare target search. Given that VGPs are known to both perform better on tasks of visual attention (e.g., Green & Bavelier, 2003) and perform faster in visual search (e.g., Castel et al., 2005), the critical goal is to determine how VGPs will perform on a visual search task in which faster speeds directly predict lower accuracy (e.g., Fleck & Mitroff, 2007; Wolfe et al., 2005).

1.3.3 Extra targets in the search environment

Visual search has typically been studied with only a single target within the distractor array. Yet, many real-world searches involve searching for more than one target at time. In radiology, many different abnormalities might be present simultaneously in the same image, and in airport security there might be more than one

hazardous item in a particular piece of luggage. How do *multiple targets* influence search behavior?

1.3.3.1 Radiology and the satisfaction of search effect

Multiple target search has been most prominently explored within the domain of radiology (e.g., Ashman, Yu, & Wolfman, 2000; Berbaum et al., 1998; Franken et al., 1994; Samuel, Kundel, Nodine, & Toto, 1995). The topic has received much attention in medical imaging because radiological research has consistently demonstrated that detecting a target in a radiographic image significantly reduces the likelihood that additional targets will also be detected in the same image or case (e.g., Berbaum et al., 1998; Franken et al., 1994). The problem has been characterized as a consequence of reaching a satisfactory interpretation of a case once a target has been found (Tuddenham, 1962), thus giving rise to the term “satisfaction of search” (SOS). SOS errors are reported across several branches of radiology, including chest radiography (e.g., Berbaum et al., 1998), abdominal radiography (e.g., Franken et al., 1994), and osteoradiology (e.g., Ashman et al., 2000).

1.3.3.2 Generality of SOS effect?

Despite evidence for SOS across multiple radiological branches and in other medical imaging searches as well (DeMay, 1997), there has been little research on the generality of the effect outside of the medical domain. However, if the effect represents a basic search heuristic common to all kinds of visual searches, there may be serious

implications for other real-world tasks. For example, airport security searches now require screeners to scan for a set of additional targets formerly allowed on flights, including liquids, gels, and toothpaste.

Importantly, these new targets are both much more prevalent and much more salient than potentially concealed targets like guns and bombs. If the SOS effect indeed generalizes to other kinds of real-world searches, it may be possible that detection of rare or hidden targets (e.g., box cutters) may be hindered by detection of easy and salient targets (e.g., a water bottle). Chapter 4 presents research that illustrates the generality of SOS as well as the various pressures which interact to modulate the effect.

Chapter 2. Expectation and rare target search

Whether looking for car keys on a desk or a friend in a crowd, we constantly engage in visual searches of our environment. Ironically, some of the most critical searches often exhibit disturbingly high rates of error: Thirty percent of malignancies are missed in radiological examinations (Berlin, 1994; Renfrew et al., 1992) and a significant percentage of dangerous items are reportedly missed in airport baggage screening. Radiology and airport screening are alike in that the targets of the search are quite rare, and a recent laboratory study (Wolfe et al., 2005) suggested that low target prevalence, per se, might directly underlie the high error rates. When searching arrays somewhat similar to those viewed by airport baggage screeners, observers missed only 7% of the targets when target frequency was high (present on 50% of trials) but an alarming 30% when target frequency was low (1% prevalence).

What drives this potentially dangerous low-prevalence effect? Wolfe and colleagues (2005) proposed that as observers repeatedly respond with correct rejections (accurately reporting that no target is present on target-absent trials), they begin to terminate their searches faster and faster, leading to misses on the rare trials that actually do contain a target. Critically though, what is the fate of the target information on those miss trials? Are observers completely unaware of the target, suggesting that they process missed targets the same as correct rejections? Or rather does the high miss rate arise from a response execution problem wherein observers actually detect the

targets, but respond too quickly? In other words, do high miss rates in low-prevalence visual search represent an error of perception or an error of action? This fundamental distinction highlights the importance of the current research project: Here we explore the origin of the low-prevalence effect with the direct goal of determining how to eliminate it.

To test whether execution errors account for the increase in misses for rare targets, we administered a similar experimental design to Wolfe et al. (2005), with the critical addition of providing observers with the opportunity to correct a previous response. Observers can presumably correct their action-based errors but not their perception-based errors, so with this simple modification we can determine if low target prevalence continues to underlie high miss rates when action errors are largely eliminated.

2.1 Methods

2.1.1 Participants and stimuli

Twenty young adults (average age=21-years-old, $SD=4.5$ years) were recruited from the Duke University community to participate in the experiment in exchange for \$15 or for course credit. All observers gave informed consent prior to participation.

The experiment was conducted on a Dell Optiplex computer running Windows 2000 and programmed in Matlab 6.5 using the Psychophysics toolbox (Brainard, 1997). Each trial began with a cross ($1.3^\circ \times 1.3^\circ$) appearing for 0.5 seconds at the center of the

screen to indicate the pending onset of the next display. The cross was replaced by the search array, which consisted of 3, 6, 12, or 18 items (see Figure 3). Objects in the search array were drawn from 30 photorealistic objects from the Hemera Photo-Objects Collections and belonged to one of five categories: *toys*, *fruits & vegetables*, *clothing*, *birds*, and *tools*. Each object was converted to grayscale and partially blurred, then presented with a random rotation in a non-overlapping array on a white background. The array of possible locations was specified by an invisible 5x5 grid (subtending $19.1^\circ \times 19.1^\circ$ at an approximate viewing distance of 60 cm), and each item (subtending on average $3.2^\circ \times 3.2^\circ$) was placed with slight spatial jitter within a randomly-selected cell, center cell excluded. On target-present trials one of the items was randomly selected from the *tool* category (e.g., hammer, wrench, clamp, saw, drill, axe) and the remaining items were drawn randomly, without replacement, from the non-*tool* categories. On target-absent trials, all items were drawn from the non-*tool* categories.

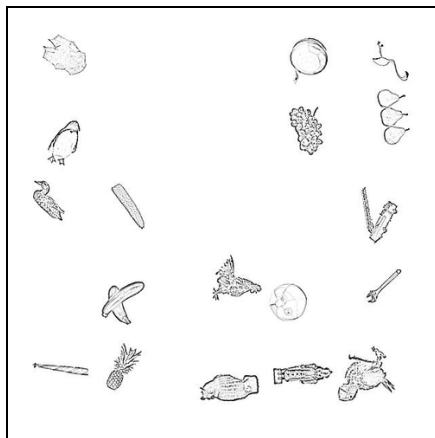


Figure 3: Sample search array. Observers searched for a tool amidst randomly rotated items from other categories.

2.1.2 Procedure

Observers searched the display for a *tool* for as long as desired, self-terminating the trial by either pressing the '/' key to indicate "target-present" or the 'Z' key to indicate "target-absent." Observers were encouraged to treat the experiment as they might an airport security task: important to keep the trials progressing, but also imperative that no "dangerous" items are missed. Upon response, the display would disappear and the next trial would automatically appear after a 0.5 second delay. Half of the observers (Correction condition) were given the opportunity to correct their response to the previous trial and half of the observers had no such option (No-Correction condition). Observers in the Correction condition were instructed to press the 'esc' key during the subsequent trial to indicate if the response to the previous trial should be reversed. No feedback was given for a response, nor was there any feedback provided after pressing the correction key—observers were told in advance that the correction would be recorded and that they should then respond to the next trial normally.

The experiment consisted of 1400 trials divided into 3 blocks by the frequency with which target *tool* items would be present. The High prevalence block consisted of 200 trials, with 50% target-present trials. The Medium prevalence block consisted of 200 trials, with 10% target-present, and the Low prevalence block consisted of 1000 trials, 2% of which contained a target. Observers were warned that most of the trials would be

very low in target frequency, and that they should resist any tendency to fall into an automatic response mode of “target-absent.” Half of the observers in each condition viewed the blocks in the order [High, Medium, Low], the other half viewed blocks in the order [Low, Medium, High]. There was no systematic effect of order for either condition and all analyses are collapsed over order. After every 200 trials, the program would prompt observers to take a break, and the experiment continued when a button was pressed. Each set of 200 trials was preceded by an onscreen indication of the target prevalence (High, Medium, or Low) of the upcoming set. Observers were strongly encouraged to take advantage of the breaks, particularly if feeling tired or bored. The entire experiment ran approximately 80 minutes in length depending on the speed of the observer.

2.2 Results and Discussion

The No-Correction condition replicated the prevalence effects of Wolfe et al. (2005) with miss rates of 10%, 19%, and 31% for the High, Medium, and Low prevalence blocks, respectively ($F(2,18) = 15.478, p < 0.001, \eta_p^2=0.632$; Figure 4). In contrast, the Correction condition showed no effect of prevalence with miss rates of 4%, 10%, and 10% for the High, Medium, and Low prevalence blocks ($F(2,18) = 1.618, p = 0.226, \eta_p^2=0.152$). A mixed-effects ANOVA revealed a significant interaction between prevalence and correction ($F(2,36) = 4.736, p = 0.015, \eta_p^2=0.208$), indicating that the prevalence-linked increase in misses specifically occurs when observers cannot correct

their mistakes. Further, the miss rates in the Correction condition calculated *before* incorporating the correction responses (8%, 19%, and 27%, respectively) were statistically equivalent to the No-Correction condition: A mixed-effects ANOVA revealed no interaction between prevalence and condition ($F(2,36) = 0.186, p = 0.831, \eta_p^2=0.010$), again highlighting the specific impact of allowing observers the opportunity to catch their own mistakes. Average false alarms rates on target-absent trials were very low for all blocks in both conditions; the High, Medium, and Low blocks in the No-Correction condition yielded false alarm rates of 0.70%, 0.22%, and 0.06%, respectively, and the Correction condition produced rates of 0.80%, 0.00%, and 0.03% respectively. The correction key was used almost exclusively to correct misses (94.4% of all corrections).

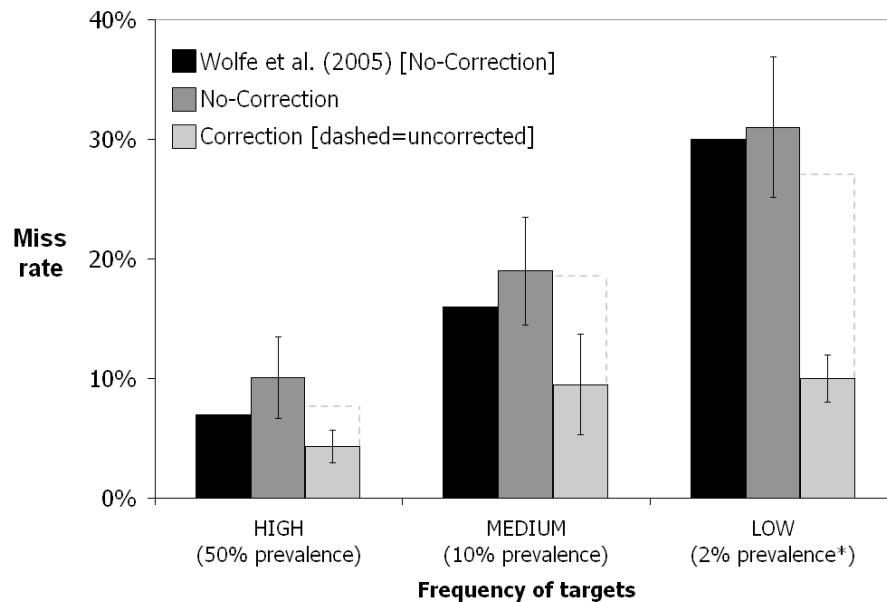


Figure 4: Miss rates by target prevalence. *Low-prevalence signifies 1% for Wolfe et al. and 2% for the current experiment.

Observers were free to respond at their own pace and their response time data are highly informative. Figure 5 shows the average response time patterns in the No-Correction condition for trials immediately before and after target trials in the Low prevalence condition, plotted separately for hits and misses. Response times for trials leading up to misses were on average 231 ms faster than trials preceding hits, in accordance with Wolfe et al. (2005) and in support of the notion that increased speeds lead to misses (Chun & Wolfe, 1996; Rabbitt, 1966). The response time data thus seem to show a direct relationship between accuracy and speed for this visual search task, and indeed, Wolfe et al. (2005) proposed that “misses seem to occur because the observers are abandoning search too quickly.” This relationship fits our proposal that searches aborted more and more rapidly with a “target-absent” response will lead to a high number of misses due to motor errors.

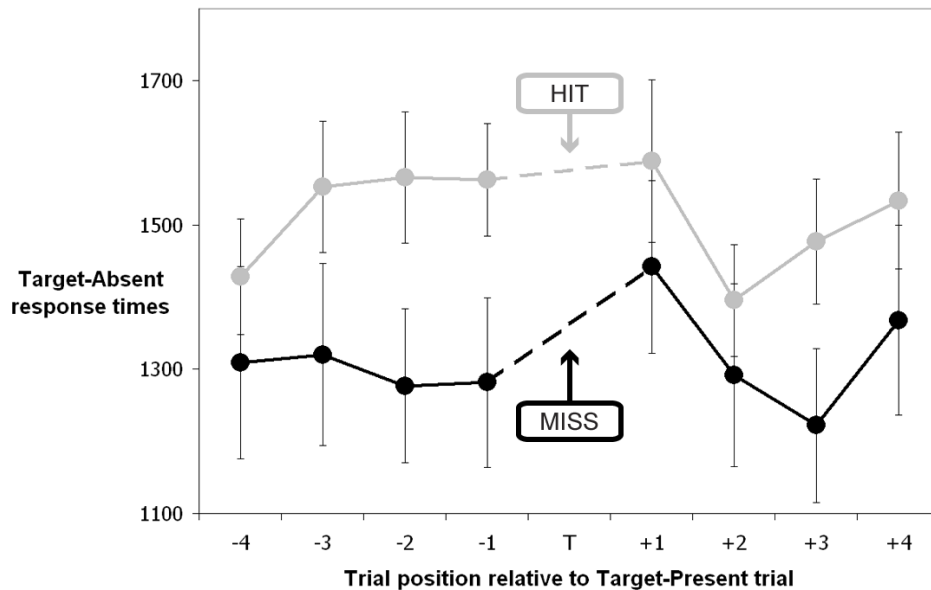


Figure 5: Low prevalence response time data from the No-Correction condition organized by ordinal relationship to a target-present trial.

However, this speed/accuracy trade-off in rare visual search has recently been challenged. New data (Wolfe et al., 2007) suggest that when observers are given “speeding tickets” on fast trials to induce slower responding overall, miss rates nevertheless remain relatively high. A direct link between response time and errors would predict improved accuracy and there are a few issues which might address this discrepancy. First, the trial duration in the speeding ticket experiment was still yoked to response speeds, rather than being fixed, which may have limited any delay-driven benefits by adding a dual task – since observers must monitor their response speed to avoid penalty, they might be judging duration while simultaneously trying to complete the search. Second, providing differential feedback for specific durations (punishment on very fast responses and nothing on slower responses) could encourage observers to

adopt the strategy of delaying the initiation of their search (and thus their response) so as to avoid penalty. Such a process by which a response rule is rescaled in reference to temporal regularities is similar to those formalized in information processing models of interval-timing (see MacDonald and Meck, 2004 for a review). Finally, it is entirely possible that the induced slowdown was simply not enough time to overcome the prepotent “target-absent” response. While these new speeding ticket results are intriguing, future work will be needed to reconcile them with the current results and an accumulating body of data (e.g., Chun & Wolfe, 1996; Wolfe et al., 2005; Wolfe et al., 2007) that have consistently revealed a relationship between faster responses and lower detection rates.

Another interesting pattern in the response time data is evident in Figure 5: Observers were on average 161 ms slower to respond on trials immediately following a missed rare target – a similar effect as that found in Wolfe et al. (2005) yet here, critically, there was no feedback provided. This slowdown joins the correction accuracy data in strongly suggesting that some missed trials actually involved cognizance of the mistake (Rabbitt, 1966) and may be processed in a similar manner as correct responses (Egeth & Smith, 1967), thus revealing an action error rather than a perceptual error. It should be noted that our miss rate and response time data fully replicate those of Wolfe et al. (2005), alleviating concerns over any methodological differences (e.g., the presence or absence of feedback).

When given the opportunity, observers readily correct their missed responses, drastically reducing the effect of target prevalence in visual search. These findings indicate that the high miss rates in the No-Correction condition largely arose from execution errors, wherein observers in fact noticed a target but responded too quickly. Whether such late recognition is driven by a lingering sensory representation of the display or whether it reflects an inability to inhibit the repetitive and prepotent “target-absent” response, it is clear that the rise in errors associated with prevalence is driven by a paradigm-specific deficit of response execution, rather than a more general perceptual issue of target identification or search failure. Since a primary aim of the present work is to relate visual search results to socially-important situations, this redefining of the low-prevalence effect is critical and demands a comparison between the response parameters of laboratory tasks and those of radiological and airport screenings.

In radiology, image readers typically spend 30 to 90 seconds on an X-ray scan and assess fewer than 100 images total in a day, a sharp contrast to the present study which had 1400 trials and average response times of less than 3 seconds. As such, an increase in misses caused by rapid responding in low-target-frequency tasks does not likely occur in this medical context, and a recent, comprehensive radiological study (Gur et al., 2003) accordingly reported no significant effects of prevalence on disease detection performance. Although target frequency is low, there are likely other mechanisms at work that might explain the high incidence of error, including interpretation deficits

(Manning, Ethell, & Donovan, 2004), “satisfaction-of-search” issues (Wolfe et al., 2005; Samuel et al., 1995), and incomplete visual scan patterns (Kundel, Nodine, & Carmody, 1978).

In contrast to radiological screening, airport baggage screening has relatively fast response times (average inspection times are 3-5 seconds; Schwaninger, Hardmeier, & Hofer, 2005) and the number of bags screened in a single session can be quite extensive. However, a direct link between baggage screening and our task remains tenuous given the differences in response parameters, stimuli, and motivation. Nevertheless, our results underscore the necessity of being able to immediately correct errors (e.g., rewind the baggage conveyor belt or more closely examine individual images) in any fast low-prevalence search. More generally, our results propose focusing less on the effect of prevalence and more on other issues that have been shown to drive high error rates in airport searches, including bag complexity, non-prototypical views of prohibited items, and overlapping x-ray images (Schwaninger et al., 2005), as well as observer-specific factors such as the ability to generalize recognition training to a diverse set of possible threat items (McCarley, Kramer, Wickens, Vidoni, & Boot, 2004).

There remains the possibility that low prevalence may interact with other factors to increase error rates in a manner yet unrevealed. While our data demonstrate no such interaction, they do emphasize that looking for any influence of prevalence need be explored when motor errors are minimized or eliminated. However, if further research

continues to establish the absence of a prevalence effect in correctable searches, this could in fact facilitate the study of misses in rare target searches (Gur, Rockette, Warfel, Lacomis, & Fuhrman, 2003; Obuchowski, 2005): Since the total number of trials needed to implement rare target searches in the laboratory can be extremely cumbersome, the present results indicate that experimenters might be able to safely inflate the number of target-present trials to better explore what mechanisms underlie high miss rates.

In sum, prevalence may not catastrophically influence error rate in correctable searches. The option to correct mistakes parses out response execution errors, thus drastically reducing the rise in miss rates previously found in rare target search. Ultimately, improving real-world search performance will best be served by separately addressing errors of action and errors of perception.

2.3 Additional findings in rare target search

The research reported in section 2.2 spawned several studies which have furthered rare target research and merit discussion here.

2.3.1 Speed/accuracy tradeoff or criterion shift?

Wolfe and colleagues (2007) utilized signal detection theory (SDT; Green & Swets, 1966) in a series of experiments to model the effects of rare targets on visual search. The authors explored the prevalence effect by using paired observers (two searchers looking at the same display simultaneously), forced slowdowns, mixing of common and uncommon target types, intermittent bursts of targets, and a “retraining”

condition of selective feedback. Although a correction option was not utilized in any of the experiments (e.g., Fleck & Mitroff, 2007), a prevalence effect was reported across a number of manipulations. The authors emphasized two main conclusions from the data: 1) observers are not simply “sloppier” under low prevalence conditions, and 2) target rarity induces a “criterion shift” whereby observers require more evidence to endorse a target as present. These interpretations speak against a simple speed-accuracy tradeoff from the point of view that an overall sloppier approach, characterized in SDT terms as reduced sensitivity, would be reflected by an increase in both missed targets and in false alarms (claiming a target is present when it was not).

However, although a criterion shift account of low prevalence fits with a similar modeling of errors in vigilance studies (e.g., Broadbent & Gregory, 1963; 1965), caution should be used in interpreting a strict SDT approach to rare target search. SDT demands that accuracy be stable and estimable across an entire experiment, yet if accuracy decreases over the course of an experiment, it becomes difficult to distinguish this change in accuracy from a shift in decision criterion (see Pastore, Crawley, Berens, & Skelly, 2003; Wixted & Stretch, 2000). A low prevalence search necessitates the repetitive responding of “target absent,” which opens up the possibility that observers’ responses become automatic, effectively eliminating the discrimination that is being measured by SDT analysis.

Another pitfall in relying too heavily on SDT theory for interpreting the prevalence effect is the extreme imbalance in trial types. Wolfe et al. (2007) base much of the criterion-shift theory on the fact that false alarms decrease significantly in the low prevalence condition. The authors suggest that global “sloppiness” should both increase misses and false alarms. However, when the majority of trials are of one particular type (target absent), a shift to an automated response of “target absent” necessarily leads to a reduction in false alarms. As a hypothetical example, imagine performing a high prevalence set of trials normally, then on low prevalence trials simply pressing “target absent” on every single trial. The net effect of reducing prevalence is an increase in miss errors, but a drastic reduction (elimination) of false alarms which theoretically indicates a criterion shift. Given an extremely consistent reduction in response times as a main effect of prevalence (Fleck & Mitroff, 2007; Wolfe et al., 2005; 2007), it is not difficult to imagine that this may be driven by participants automatically responding on some or many trials as a function of boredom or fatigue.

Although the data presented in the previous section indicated that low prevalence induces speed-related, correctable motor errors, it is important to acknowledge that under similar circumstances, observer may adopt speeds at which errors are indeed non-correctable: at high speed, not all targets will be fixated. Thus, a low prevalence effect exists, but the main point is that the effect is strongly driven by the change in speed. One condition in Wolfe et al. (2007) offers evidence contrary to this

position, wherein observers were induced to slow down in an attempt to minimize error. Yet, observers continued to exhibit a low prevalence effect. This data is difficult to reconcile with a straight speed-accuracy tradeoff interpretation, but see the discussion in section 2.2 for a proposed dual-task explanation involving how observers were slowed down.

Finally, postulating a speed-accuracy tradeoff as the primary mechanism of low prevalence errors does not preclude a simultaneous criterion shift as described in Wolfe et al. (2007), and indeed the data from early vigilance studies (e.g., Broadbent & Gregory, 1963; 1965) may indeed be applicable here. However, it may be extremely difficult to deconvolve the two contributions without explicitly modeling decision parameters (see Madden et al., in revision). Nevertheless, the consistent finding that miss trials at low prevalence are terminated much more quickly than the time necessary to find a target suggests that a speed-accuracy tradeoff may be the more parsimonious explanation.

2.3.2 Change in quitting threshold

Rich et al. (2007) also explored the relative contributions of different sources to the low prevalence effect, in part by tracking eye movements during a rare target search experiment. The authors demonstrated the prevalence effect in both trivial feature searches (e.g., find a red square amongst blue squares) and more complicated configuration displays (e.g., find a red square amongst red circles and blue squares).

However, eye position data suggested different primary sources of errors: Configuration search errors derived primarily from a change in quitting threshold (i.e., terminating a search earlier), whereas feature search generated motor errors (e.g., Fleck & Mitroff, 2007). The quitting threshold hypothesis has been discussed previously in terms of how to model the process by which a searcher terminates a search (Chun & Wolfe, 1996). Simply, the claim is that observers will speed up after correct responses, and slow down after incorrect responses.

Rich and colleagues (2007) make no explicit claims about a criterion shift, instead focusing on distinguishing between the contributions of a changing quitting threshold and motor errors. Yet, it may in fact be parsimonious to consider motor errors as indeed a manifestation of a shifted quitting threshold. Presumably, if observers change quitting thresholds as a function of target rarity, at some point the speed should be fast enough to induce motor errors by not spending enough time to inhibit a prepotent response. While certain conditions may certainly facilitate this particular situation, such as the featural searches demonstrated by Rich et al. (2007), all errors essentially arise when searches are terminated too quickly.

Furthermore, we maintain that the quitting threshold hypothesis directly maps onto the speed-accuracy tradeoff account supported by Fleck & Mitroff (2007). As target prevalence decreases, responses to trials speed up. As a consequence, the speeded

decisions (i.e., faster quitting thresholds) lead to less discrimination, and this is likely the primary source of rare target error.

2.3.3 Correction does not always eliminate the prevalence effect

Van Wert et al. (2009) have recently conducted a series of studies replicating the correction option of Fleck & Mitroff (2007). The studies utilized realistic x-ray images of luggage and manipulated the feedback given to observers. Errors were corrected either with a follow-up button press (as in Fleck & Mitroff, 2007) or with a required secondary response to confirm the initial response. The authors continued to report a significant prevalence effect, although they did report that the absence of feedback tended to reduce low prevalence errors. As in Wolfe et al. (2007), the effects were attributed again to shifts in decision criteria. The authors conclude that the simple stimuli used in Fleck & Mitroff (2007) may account for the reduced prevalence effect associated with the option to correct, and that a true prevalence effect is more clearly associated with the complex stimuli of these experiments. Although Van Wert et al. (2009) attribute the difference in findings to the complexity of stimuli, it is worth noting several other differences in both paradigm and results that may be relevant.

First, unlike Fleck & Mitroff (2007) and Wolfe et al. (2007), the false alarm rate did not significantly decrease between high and low prevalence. The authors attribute this lack of change in false alarm rate to the absence of feedback, yet it should be noted that Fleck & Mitroff (2007) as well eliminated feedback and also showed the typical

finding of more false alarms at high prevalence (3 trials out of 100) than at low prevalence (less than 1 trial out of 100; see section 2.2). Importantly, since a decline in false alarm rate was cited as evidence against a speed-accuracy tradeoff in previous work from this lab (Wolfe et al., 2007; see section 2.3.1), the absence of such an effect here suggests a cautious interpretation of the proposed criterion shift account.

Next, Van Wert et al. (2009) showed no effect of prevalence on response times, again in sharp contrast to most other studies of prevalence (Fleck & Mitroff, 2007; Rich et al., 2007; Wolfe et al., 2005; 2007). No explanation of the difference is offered in the paper, but such a finding would speak against a speed-accuracy tradeoff account of the prevalence effect. However, the second experiment in Van Wert et al. (2009) required observers to make a second response and to spend more time on the trial, which led to an improvement in accuracy. Although a prevalence effect still remained in the response confirmation condition, it is notable that extra time spent searching increased overall accuracy and implicates some degree of a speed-accuracy tradeoff.

Lastly, Van Wert et al. (2009) suggest that the prevalence results from Fleck & Mitroff (2007) are mostly driven by differences in stimulus complexity. However, one other critical difference should be noted: the average age of participants. In Fleck & Mitroff (2007), participants were primarily Duke undergraduate students (average age=21 years), whereas in Van Wert et al. (2009), the participants were local volunteers typically many years out of college (average age = 27, 32, and 34 across three studies).

While neither a criterion shift nor a speed-accuracy tradeoff predicts a specific difference between age groups, age may be a factor to consider when understanding the numerous differences in findings between Van Wert et al. (2009) and previous work.

Recent data has explored the effects of aging on rare target search (Madden et al., in revision), implicating a compensatory shift in response strategy and behavior to accurately respond to rare targets. Although that data explored young versus elderly effects, there may be similar differences in strategy or motivation between college-aged and 30-year-old participants. Younger participants may have extensive video game experience, for example, potentially offering a strategic set of guidelines in performing the task. At the same time, a 19-year-old student walking in to participate between classes for a required credit may have significantly different motivation than an individual 10 years older, driving in to participate for the sake of science and monetary compensation. The effects of video game experience and motivation are extensively discussed in the next chapter.

Chapter 3. Video game experience improves rare target search

Playing fast-paced action video games results in substantial changes in performance on visual attention and perception tasks (e.g., Dorval & Pepin, 1986; Drew & Waters, 1986; Gagnon, 1985; Green & Bavelier, 2003; Greenfield, DeWinstanley, Kilpatrick, & Kaye, 1994; Griffith, Voloschin, Gibb, & Bailey, 1983; Lintern & Kennedy, 1984; Subrahmanyam & Greenfield, 1994). Expert video game players (VGPs) exhibit enhanced attentional abilities over non-video game players (NVGPs) in both spatial and temporal domains: Compared to NVGPs, VGPs can track more moving objects, (Green & Bavelier, 2006b; Trick, Jaspers-Fayer, & Sethi, 2005), have a wider spatial distribution of attention, even at distances outside the typical video game field of view (Green & Bavelier, 2006a), and miss fewer targets in a rapidly presented stream of visual stimuli, suggesting a higher temporal resolution of attention (Green & Bavelier, 2003). Across many tasks, VGPs also respond significantly faster than NVGPs (e.g., Castel et al., 2005). VGP-like benefits are observed in NVGPs after video game exposure, suggesting that video game effects on attention and perception are not simply the result of self-selection (e.g., De Lisi & Cammarano, 1996; Gopher et al., 1994; Green & Bavelier, 2003, 2006a,b,c, 2007; but see Boot et al., 2006; Gagnon, 1985; Rosenberg et al., 2005; Sims & Mayer, 2002).

The effects of video game experience on visual attention and perception have been characterized as a *low-level* modulation of visual processing (e.g., Green & Bavelier,

2003, 2007) in which video game experience creates an enhanced attentional capacity, allocation of spatial attention, and ability to process information over time (Green & Bavelier, 2003). Yet, such hypothesized low-level changes to the visual system may be confounded with *higher-level* cognitive differences between VGPs and NVGPs. If VGPs were to possess relatively heightened motivation or enhanced strategies, should video game related effects be characterized as low-level changes to perceptual processing, as higher-level differences in approaches to tasks, or both?

Low-level perceptual changes due to specific experiences have been explored within the domain of perceptual learning. Indeed, an analogous query to that posed here has been asked by perceptual learning theorists (Goldstone, 1998): To what degree is perceptual learning best characterized as low-level changes to the processing stream versus higher-level cognitive learning? Clear evidence for low-level based changes arose from findings that perceptual learning regimens produce physiological differences far too early in the information stream (e.g., 100ms after stimulus presentation) to ascribe to higher-level learning (Fahle & Morgan, 1996; Goldstone, 1995; Sekuler, Palmer, & Flynn, 1994). Early temporal aspects of perceptual learning and their associated physiological underpinnings underlie an extreme specificity which is considered a signature characteristic of perceptual learning; for example, learning to discriminate slight changes in the orientation of vertical lines does not transfer to horizontal lines (Fiorentini &

Berardi, 1980), and learning with one eye does not transfer to the other (Karni & Sagi, 1991).

Video game effects, in contrast, are particularly exciting precisely because of a lack of specificity; VGPs show enhanced performance on a wide variety of tasks that assess a diverse array of abilities and even excel on tasks outside typical video game fields of view (e.g., Green & Bavelier, 2003, 2006a). Just as the task-specificity of perceptual learning supported a low-level account for behavioral changes, the task-generality of video game playing effects may indicate higher-level cognitive differences that underlie VGPs' enhanced performance. While it is possible that video game effects reveal a new generalized form of low-level perceptual learning, we explore here whether, compared to NVGPs, VGPs are a) more motivated to perform well, b) more aware of their own cognitive and perceptual abilities, and c) willing to adjust their responding to improve performance. Note that similar questions have been raised before, but not explicitly tested (Green & Bavelier, 2006b,c).

No video game study has yet examined the possible role of higher-level effects on attention-demanding tasks, and to do so here we utilize a visual search task which is particularly sensitive to motivational and strategic differences. Specifically, we exploit an important discovery about the nature of visual search: Subjects consistently produce more miss errors when targets are rarely present compared to when they are frequently present, due to a combination of motor and perceptual failures (Fleck & Mitroff, 2007; Li

et al., 2007; Rich et al., 2008; Van Wert, Horowitz, & Wolfe, 2009; Wolfe et al., 2005; 2007). Low target prevalence search necessitates extended testing sessions and within such situations, participants decrease the amount of time spent on each trial (relative to high prevalence searches) and subsequently miss more rare targets. Note that this response pattern diverges from standard vigilance paradigms wherein participants increase their response timing over the course of long experimental sessions rather than decrease (Buck, 1966). A rare target search task thus offers a useful case study for examining the influences of higher-level video game effects because it is a) self-paced, allowing participants to establish their own response speeds, and b) prone to many errors, yet errors that are easily avoided if participants are willing and able to adjust their speed. Importantly, VGPs typically respond faster than NVGPs (e.g., Castel et al., 2005), yet rare target searches are particularly sensitive to speeded responses (e.g., Fleck & Mitroff, 2007; Rich et al., 2008). Will VGPs be more accurate than NVGPs in high and/or low prevalence visual search, and if so, will their response times reveal motivational influences?

3.1 Methods

3.1.1 Participants and stimuli

Twenty male participants (average age=20.8-years-old, $SD=2.8$ years; VGP=19.7-years-old and NVGP=21.9-years-old) were recruited from the Duke University community and received \$15 or course credit. After the experiment, participants

completed a questionnaire which assessed their amount of video game experience and recency of play for 6 video game genres. VGP status was determined by a weighted scoring of the questionnaire with an emphasis placed on extensive and recent experiences with fast-paced action games (e.g., NCAA basketball, Madden football) and “first person shooters” (e.g., Medal of Honor, Gears of War, Halo). Many of the participants completed a video game questionnaire prior to the day of their experimental session as part of a larger, unrelated recruiting effort and these preliminary responses were used as a means to selectively recruit likely VGPs and NVGPs. As a result a split of the participants (10 in each group) resulted in VGPs having both recent and extensive first-person shooter experience and NVGPs having very little or no such experience. Participants were not informed that the study was concerned with video game experience until after they had completed the experiment. Three additional participants (2 VGPs, 1 NVGP) were excluded from analyses as outliers for both response time and accuracy.

The experiment was conducted on a Dell Optiplex computer running Windows 2000 and programmed in Matlab 6.5 using the Psychophysics toolbox (Brainard, 1997). Objects in the search array were drawn from 30 photorealistic objects from the Hemera Photo-Objects Collections and belonged to one of five categories: *toys, fruits & vegetables, clothing, birds, and tools*. Each object was converted to grayscale and partially blurred,

then presented with a random rotation in a non-overlapping array on a white background (see Fleck & Mitroff, 2007).

3.1.2 Procedure

The experimental methods were identical to the correction condition of Fleck & Mitroff (2007). Each VGP and NVGP performed a basic visual search task wherein participants made a self-paced “target-present” or “target-absent” response. Each trial consisted of 3, 6, 12, or 18 items. On target-present trials one of the items was randomly selected from the *tool* category and the remaining items were drawn randomly, from the *non-tool* categories. On target-absent trials, all items were drawn from the *non-tool* categories. Participants self-terminated each trial by either pressing the ‘/’ key to indicate “target-present” or the ‘Z’ key to indicate “target-absent.” Additionally, all participants were given the option to correct mistakes by pressing a third button (the ‘esc’ key) during the subsequent trial. This keypress would indicate that the response to the previous trial should be reversed, but participants were not allowed to go back and look again to make sure. No feedback was given at any point.

Participants received printed and verbal instructions and were shown un-rotated images of the six possible target *tools*. Participants were encouraged to treat the experiment as they might an airport security task: important to keep the task progressing, but also imperative that no “dangerous” items are missed. The experiment began with 50 practice trials, followed by 1400 test trials divided into 3 blocks of

differing target prevalence (i.e., frequency with which a target *tool* item would appear). The High prevalence block consisted of 200 trials, with 50% target-present. The Medium prevalence block consisted of 200 trials, with 10% target-present, and the Low prevalence block consisted of 1000 trials, 2% of which contained a target. Participants were warned that most of the trials would be very low in target prevalence, and that they should resist the tendency to fall into an automatic response mode of “target-absent.” Half of the participants in each group performed the block order of [High-Medium-Low], the other half received [Low-Medium-High]. Participants were prompted to take a break every 200 trials and during each break were informed of the target prevalence (High, Medium, or Low) of the upcoming trial set.

3.2 Results

Trials with durations shorter than 300ms or greater than 10000ms were excluded from analyses, which eliminated 2.3% of trials for the 20 participants.

3.2.1 Visual search accuracy

A repeated measures ANOVA on mean accuracy rates was conducted with Video Game Experience (VGP or NVGP), Correction (including or excluding made corrections), and Block Order (Low-Medium-High or High-Medium-Low) as between-subjects variables, and Target Prevalence (Low, Medium, High) and Set Size (3, 6, 12, 18) as within-subjects variables. There were main effects of correction ($F(1,16)=34.32, p<.001$), prevalence ($F(2,32)=14.88, p<.001$), set size ($F(3,48)=2.83, p=.048$), and video game

experience ($F(1,16)=6.63, p=.020$), but not block order ($F(1,16)<1, p>.05$). This reveals that the error rate decreased when 1) participants used the correction option, 2) prevalence increased, and 3) set size decreased.

There was a significant interaction between correction and video game experience ($F(1,16)=6.63, p=.020$), revealing that NVGPs corrected more errors than VGP, although this is likely driven by VGPs making few mistakes that needed to be corrected. There was a significant interaction between prevalence and video game experience ($F(2,32)=8.26, p=.001$), revealing that VGPs missed far fewer targets than NVGPs specifically at the Low prevalence level (VGP $M=4.8\%$, $SD=9.2\%$; NVGP $M=17.9\%$, $SD=9.5\%$). This difference was less pronounced at the Medium prevalence level (VGP $M=5.5\%$, $SD=5.7\%$; NVGP $M=7.2\%$, $SD=5.7\%$) and High prevalence level (VGP $M=1.6\%$, $SD=1.9\%$; NVGP $M=3.5\%$, $SD=1.9\%$). Additional interactions unrelated to our hypotheses are not reported here. False positive rates were extremely low in all cells, precluding a meaningful signal detection analysis.

Real-world rare target search tasks (e.g., airport security screening and radiology) often allow searchers an option to correct self-realized mistakes. To relate the current data to such searches, we conducted a planned t-test on the post-correction accuracy data for VGPs and NVGPs within the Low prevalence condition, revealing a significant effect of video game experience ($t(18)=2.31, p=.033$). VGP experience

predicted fewer missed rare targets ($M=3.5\%$, $SD=6.9\%$) than NVGP experience ($M=10.5\%$, $SD=6.9\%$) (Figure 6).

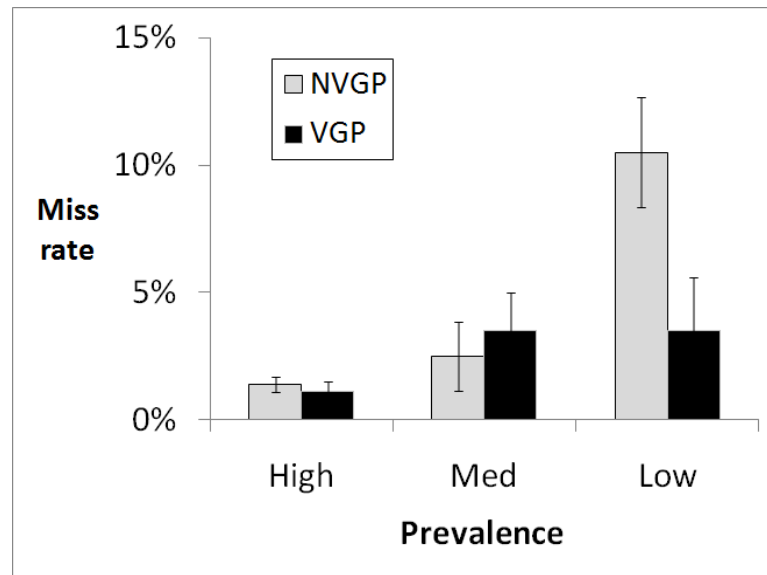


Figure 6: Post-correction miss rates for low-prevalence trials for VGPs and NVGPs at each Prevalence level.

3.2.2 Response times

We conducted a repeated measures ANOVA on response times with Video Game Experience (VGP or NVGP) and Block Order (Low-Medium-High or High-Medium-Low) as between-subjects variables, and Target Prevalence (Low, Medium, High) and Set Size (3, 6, 12, 18) as within-subjects variables. Of particular interest to our hypothesis, there was a marginally significant interaction between prevalence and video game experience ($F(2,32)=3.18$, $p=.074$). Examination of response times collapsed across set size and block order (Figure 7) revealed that VGPs systematically slow their responses as prevalence decreases (High: $M=1497ms$, $SD=288ms$; Medium: $M=1616ms$,

$SD=376ms$; Low: $M=1704ms$, $SD=503ms$), whereas NVGPs change very little across prevalence levels (High: $M=1471$, $SD=291ms$; Medium: $M=1419ms$, $SD=386ms$; Low: $M=1416ms$, $SD=512ms$). There was one main effect: responses times increased as a function of increasing set size ($F(3,48)=202.34$, $p<.001$). Additional interactions unrelated to our hypotheses are not reported here.

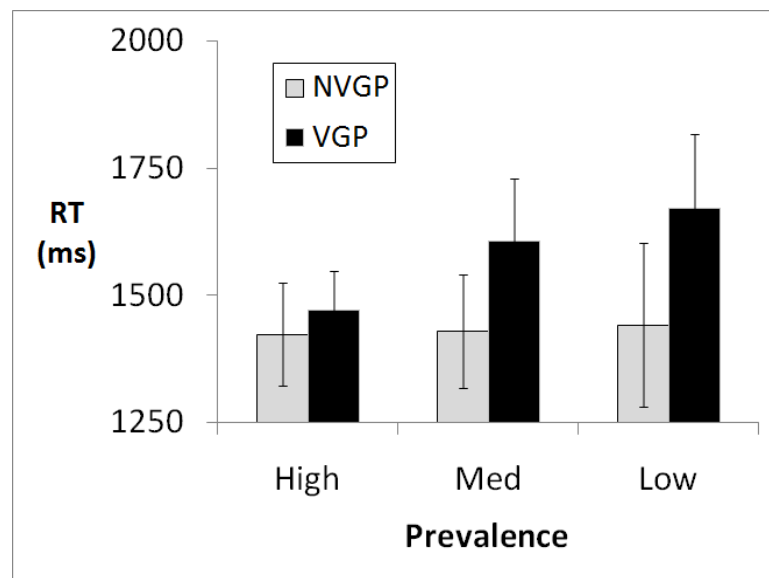


Figure 7: Response times for low-prevalence trials for VGPs and NVGPs at each Prevalence level.

3.2.3 Speed-accuracy tradeoffs

Together, the accuracy and response time data indicate that VGPs are both more accurate at low prevalence search and take more time at completing the task. This relationship between speed and accuracy can be observed directly when plotting response times and miss rates for each individual participant at each prevalence level (Figures 8a-c). These plots indicate that even as prevalence decreases, and task difficulty

accordingly increases, NVGPs tend to remain clustered around a similar response time as other NVGPs (~1450ms) even as errors increase, whereas VGPs appear to take both more time at lower prevalence and show a greater spread in the range of response times.

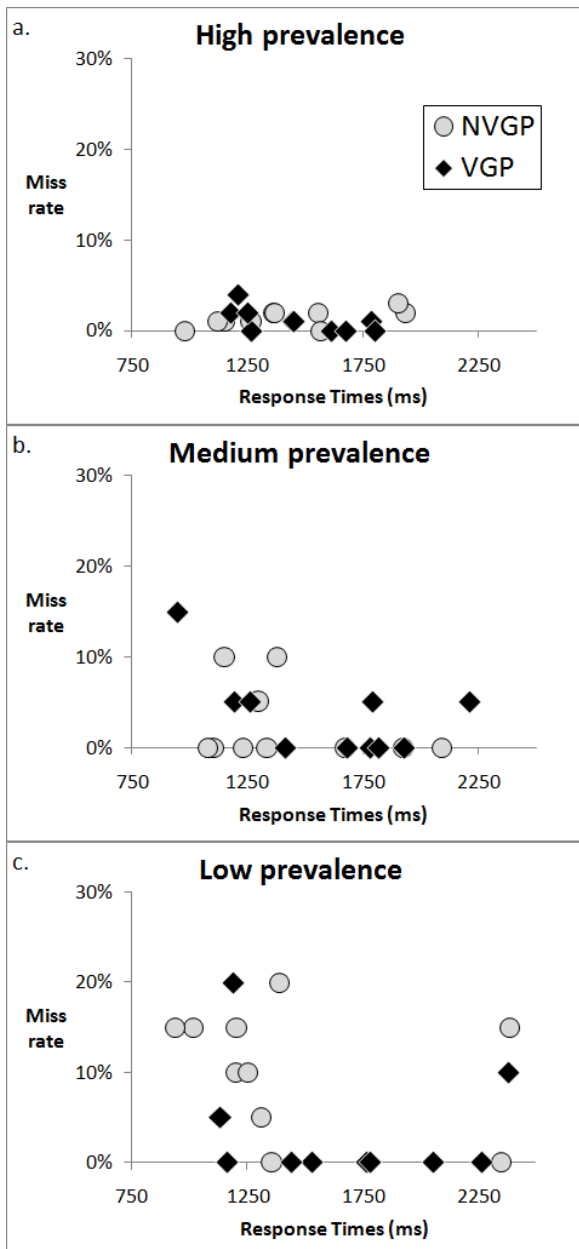


Figure 8: Speed and accuracy are plotted for each participant at (a.) High prevalence, (b.) Medium prevalence, and (c.) Low prevalence.

3.3 Discussion

3.3.1 Video gamers excel at rare target visual search

Individuals with extensive video game experience (VGPs) and with little or no experience (NVGPs) participated in a visual search paradigm designed to simulate an airport-baggage screening task. By manipulating target prevalence we examined whether prior video game experience modulates performance between a relatively easy and accurate search (high prevalence) and relatively difficult and error-prone search (low prevalence). For high prevalence search, both VGPs and NVGPs performed well and at a comparable level (c. f., Castel et al., 2005). In contrast, for low prevalence search, NVGPs missed significantly more targets than at high prevalence while VGPs remained highly accurate.

Why and how do VGPs perform so much better at rare target search than NVGPs? At a first pass, two sources of response time evidence would suggest VGPs should actually be more error prone than NVGPs. First, participants in general respond quickly under conditions of rare target search and this speeded responding results in high error rates (Fleck & Mitroff, 2007; Rich, et al., 2008; Van Wert et al., 2009; Wolfe, et al., 2005; 2007). Second, VGPs typically respond faster than NVGPs and have even been shown to do so in high prevalence search tasks (Castel et al., 2005). These factors in conjunction may a priori predict that VGPs will respond quicker than NVGPs in rare target searches and thus might miss more targets. Yet, VGPs in fact slowed their

responses within the rare target condition (while NVGPs demonstrated the typical speed-up), which directly accounts for their increased accuracy. VGPs and NVGPs revealed drastically different accuracy vs. response time relationships: VGPs were consistently at near-perfect accuracy, but demonstrated a wide range of individual response times, whereas NVGPs consistently responded at a similar speed (~1450ms) but with variable rates of accurate target detection (Figure 8c). VGPs adjusted their individual speed to maximize accuracy while NVGPs responded in a time window that produces 'good enough' accuracy.

A common question in video game research probes the causal nature of observed VGP benefits: Are VGPs better because they play video games or do they play video games because of a pre-existing benefit? Some prior studies have exposed NVGPs to video games to see if they would reveal VGP-like benefits. Several have found significant effects of the training (e.g., De Lisi & Cammarano, 1996; Gopher et al., 1994; Green & Bavelier, 2003, 2006a, 2007), but others have not (e.g., Boot et al., 2006; Gagnon, 1985; Rosenberg et al., 2005; Sims & Mayer, 2002). While future training studies may be informative with the current paradigm to see if video games can be used as a motivational tool, what is important here is that prior video game exposure reveals a systematic relationship between response speed and accuracy.

3.3.2 Role of VGP motivation in enhanced rare search performance

Previous VGP benefits have been attributed to a greater attentional capacity, better allocation of attention over space, and higher temporal resolution of attention (e.g., Green & Bavelier, 2003, 2006a, 2007), but the response time differences in the present experiment offer an important case study that suggests an alternate explanation. The VGPs' slowdown solely within the condition in which speed is directly related to errors suggests they adopt a vastly different, higher-level approach to rare target search. The VGPs' near-perfect performance in both high and low prevalence conditions, suggests they are a) meta-cognitively aware of their own limitations (e.g., "how fast can I go and still be accurate?"), and b) motivated to make the necessary behavioral adjustments to remain accurate, even at the expense of making a long experiment even longer. Two possibilities might explain the NVGPs lack of a similar slowdown on the more error prone rare-target blocks. First, they might be unaware of the increased likelihood of making errors in a rare-target situation or unaware that they are making more errors, and thus maintain similar response times for both high and low prevalence searches. Second, NVGPs might be aware of the increased difficulty and their own limitations, but simply unwilling, unmotivated, or unable to exert the extra effort to change their behavior and lengthen the task.

The significant reduction in response times for VGPs compared to NVGPs for rare-target search points to higher-level cognitive differences, above and beyond

possible low-level modulations in perceptual learning abilities. This higher-level explanation for group difference poses interesting theoretical possibilities; at an extreme, if the VGPs' rare target search accuracy benefits are solely driven by their slowdown, then NVGPs' accuracy should be able to be raised to the level of VGPs by slowing their responses, either in a forced paradigm or by reward or punishment. Similarly, forcing VGPs to respond at a faster rate should induce error rates comparable to NVGPs.

3.3.3 Implications for prior and future video game research

The present data suggest that VGPs approach difficult experimental tasks with relatively heightened motivation and/or a greater willingness to develop and apply beneficial strategies. Converging empirical evidence has previously suggested this possibility: When briefly shown arrays of items and asked to report how many were present (i.e., 'subitizing'), VGPs were more accurate and quicker for small number arrays but more accurate and slower for large number arrays (Green & Bavelier, 2006b). While this accuracy-response time pattern was interpreted as revealing a low-level modulation (e.g., a larger visual working memory capacity for VGPs), it could also be interpreted as a strategic compensation. VGPs typically respond faster than NVGPs so evidence for VGPs selectively increasing their response times in difficult scenarios nicely reveals a higher-level influence on performance. However, higher-level effects do not necessitate longer response times—motivational differences between VGPs and NVGPs

can take many forms and should be considered within the context of previously reported video game differences.

The extent to which higher-level differences drive VGPs' benefits remains an open and exciting question. On one hand, some video game effects appear primarily low-level—e.g., VGPs, and NVGPs trained on action video games, reveal enhanced visual acuity (Green & Bavelier, 2007). On the other hand, the current low prevalence search results appear primarily driven by higher-level differences. In the middle reside several findings that may benefit from both low- and higher-level effects. For example, VGPs can track more moving objects (Green & Bavelier, 2006b; Trick, et al., 2005), spread their attention across a wider field of space (Green & Bavelier, 2006a), and better detect rapid visual events (Green & Bavelier, 2003). Might such effects be, at least partially, driven by a greater motivation to perform well? Might VGPs be more willing to exert effort on computer-based laboratory tasks, treating them as another video game to master?

3.3.4 Conclusions and broader implications

The present study demonstrated that video game expertise predicts better detection of rare targets. Importantly, this performance benefit arises from VGPs drastically slowing their responses specifically during low prevalence search in which faster speeds lead to more errors. This response time choice by VGPs suggests that they prioritize accuracy more so than NVGPs, and are able and willing to adjust their speed accordingly. Such higher-level differences highlight potentially broad and generalized

motivational effects for performance. Claims of low-level perceptual differences arising from video game experience must be viewed in light of VGPs simply being willing to devote more energy, effort, and interest to the task at hand.

The current findings speak directly to an ongoing discussion about the nature of rare target search (e.g., Fleck & Mitroff, 2007; Van Wert et al., 2009; Wolfe et al., 2007); the individual difference response time and accuracy effects highlight the tradeoff of speed/accuracy to rare target error rates. Further, because rare target search serves to simulate the target rarity of several socially-critical tasks, the present evidence for effects of video game experience may also inform real-world domains with similar target frequencies, including airport luggage search (e.g., Schwaninger, 2005), routine radiological examinations (e.g., Gur et al., 2003), industrial inspections (Drury & Addison, 1973), and cytological screening (Bowditch, 1996; Wilbur, 1997). We demonstrate here dramatic effects of higher-level differences and how such differences can be selectively harnessed. This explanatory mechanism provides great promise for future study and implementation for both science and society.

Chapter 4. Extra targets in the visual search environment

Missing an abnormality in a radiological examination can have dire consequences. As such, radiological research has scrutinized the circumstances of missed targets. Among other findings, it has been shown that a specific target is more likely to be missed when it is accompanied by an additional abnormality than when it is the only target in a radiological scan (e.g., Tuddenham, 1962). That is, an abnormality has a higher rate of detection when presented alone than when presented in the same image as another problem spot. This phenomenon was originally characterized as a visual search that is discontinued once the searcher finds a target, and then becomes “satisfied” with the meaning of an image (Tuddenham, 1962). Such “satisfaction of search” (SOS) errors remain an acknowledged problem in radiologic examinations and have been demonstrated in chest radiography (e.g., Berbaum et al., 1998; Samuel et al., 1995), abdominal radiography (e.g., Franken et al., 1994), osteoradiology (e.g., Ashman et al., 2000), and multiple trauma patients (e.g., Berbaum et al., 2007).

SOS generalizes across multiple radiological domains and is known to extend, at a minimum, to certain cytological searches as well (Bowditch, 1996). But do SOS errors arise in other visual searches beyond the medical domain? The goal of the current paper is to explore the parameters in which SOS occurs in non-medical visual search in order to simultaneously inform multiple fields. First, expanding the study of SOS into non-

medical searches will contribute to the discussion about the cause of SOS within radiology. Recent radiography studies have suggested that SOS effects can arise from scanning errors (Berbaum et al., 1996; Samuel et al., 1995), recognition errors (Berbaum, Franken, Dorfman, Caldwell, & Krupinski, 2000), and/or decision errors (Franken et al., 1994). Understanding the parameters of SOS will be important in better specifying the source of such errors and developing methods to counteract them.

Second, establishing the scope of SOS errors in non-medical searches can inform the nature of the basic cognitive processes broadly involved in visual search. SOS may reflect a general heuristic of the decision-making process (e.g., “satisficing”, see Simon, 1976) involved in any kind of visual search, and if so, such errors should be observable in non-radiological contexts. Recent studies incorporating searches with more than one target (e.g., Menneer, Barrett, Phillips, Donnelly, & Cave, 2007; Wolfe et al., 2007) have complemented and expanded general theories of visual search (e.g., Treisman & Gelade, 1980; Wolfe, 2007) and the current study can contribute further. Past research has examined various external factors on search such as item familiarity (e.g., Wang, Cavanagh, & Green, 1994) and the emotional salience of search items (e.g., Hanson & Hanson, 1988), and here we explore the effects of multiple targets.

Third, finding that SOS effects exist outside the medical realm carries critical implications for real-world tasks such as airport security searches. For example, does the presence of water bottle in a luggage X-ray adversely affect the detectability of a pair of

scissors also in the bag? While the commonalities between airport baggage screening and medical image searches have only briefly been considered together (e.g., Fleck & Mitroff, 2007; Gale, Mugglestone, Purdy, & McClumpha, 2000; Wolfe et al., 2005), given the dangerous implications it is critical to determine if multiple target errors might occur in airport security searches and to establish what properties of the search might be predictive of SOS in these critical situations.

In the present paper, we explore the robustness of SOS in a controlled laboratory setting where we can simultaneously manipulate several factors. Through 13 experiments we examine how accuracy in a dual-target search is affected by (1) the relative salience and frequency of different target types, (2) stimulus discriminability, (3) time pressure, (4) perceptual set, (5) search instructions, and (6) reward pressure. We reveal that SOS errors (the reduced detection of a target when another target has also been detected in the same display) may reflect a default and generalized search heuristic that is sensitive to the interplay between several factors.

In Experiments 1-3, participants search for Ts amongst Ls and on any given trial there can be either zero, one, or two target Ts. The key manipulation across the three experiments is the relative frequency of easy-to-spot and hard-to-spot Ts; holding all else constant we vary the relative frequency of high-salience targets and low-salience targets to examine the effect of differing expectations about the presence of different target types. In Experiments 4 and 5, we explore the contribution to SOS errors of

decision-making processes (i.e., having to discern whether a fixated item is a target or distractor) by replacing the Ts and Ls stimuli with photographic images. Whereas the Ts and Ls require additional interrogation after fixation to distinguish a target from a distractor, the picture stimuli can be easily and quickly categorized. In Experiments 6-8 we examine the role of time pressure and the interplay between time pressure and the relative frequency of high- and low-salience targets, by doubling the amount of time allowed for each trial. In Experiments 9 and 10 we examine the role of “perceptual set” in driving SOS by using dual-target searches where both targets are equally salient. In Experiment 11, we examine the effect of instruction on expectation by explicitly directing participants to find high-salience targets before searching for low-salience targets. Finally, in Experiments 12 and 13, we implement a performance-based reward system to explore how pressure to maximize both accuracy and speed influences SOS errors.

4.1 Experiments 1-3: Examining the effect of salience and expectation on SOS

SOS in radiology is typically defined as an increased detection rate for a particular target (e.g., a lesion) when it is the only target in a radiographic image compared to when it is accompanied by an additional target (e.g., a pulmonary nodule). The relative salience of the primary and secondary targets is not usually manipulated in such studies and typically the targets are from different categories (e.g., a lesion and a pulmonary nodule). However, evidence from osteoradiology found that SOS was

stronger when the added secondary target was more salient than the test target (Berbaum et al., 1994). Cytology research has additionally suggested that the frequency and conspicuity of particular targets may lead to missing subtle, smaller targets (Bowditch, 1996; DeMay, 1997). Here we are particularly interested in this situation with differing salience between the primary and secondary target given that, in response to the nature of recent terrorist threats to aircraft safety, airport baggage security regulations were broadened in August 2006 by adding new categories to the prohibited items list, including liquids and gels (Transportation Security Administration website, retrieved Feb., 2009). These additions vastly increase both the frequency and salience of certain targets, as well as the likelihood of multiple targets co-occurring in the same image. Here we ask how the addition of a highly salient and frequent target (e.g., a water bottle) might impact the detection of a less salient and/or infrequent target (e.g., a box cutter) which may be present in the same image. Manipulating the frequency of a particular target type may bias the searcher's expectation for that target and contribute to producing an SOS effect.

4.1.1 Methods

4.1.1.1 Participants and stimuli

Thirty individuals (Experiment 1: mean age=18.9 years, $SD=1.2$ years; Experiment 2: mean age=18.4 years, $SD=1.0$ years; Experiment 3: mean age=18.7 years, $SD=0.5$ years) from the Duke University community participated in the study (10 in each

experiment. Each participant in this paper completed only one experiment). All participants gave informed consent and received either course credit or \$10. The experiments were conducted on a Dell Optiplex computer running Windows 2000 and programmed in Matlab 7.0 using the Psychophysics toolbox version 3.0 (Brainard, 1997).

The stimuli were comprised of two perpendicular lines slightly offset from each other (Ts and Ls, stroke width=0.3°, subtending 1.3° x 1.3° total), with target Ts having a crossbar directly in the middle and distractor Ls having the crossbar slid at variable distances away from the center. Stimuli were presented on a rendered grayscale “cloud” background (brightness range=10-50 % black) that differed on each trial (Figure 9). Distractor Ls were presented at varying shades of gray (range=28-66 % black), and target Ts were presented at one or both of two visibility levels: high-salience (range=66-70 % black) or low-salience (range=28-40 % black). In this fashion, high-salience Ts were relatively easy to detect and low-salience Ts were more difficult to detect. Each stimulus was placed with a slight spatial jitter within randomly selected cells of an invisible 8x7 grid subtending 25.4° x 19.1° at an approximate viewing distance of 60 cm.

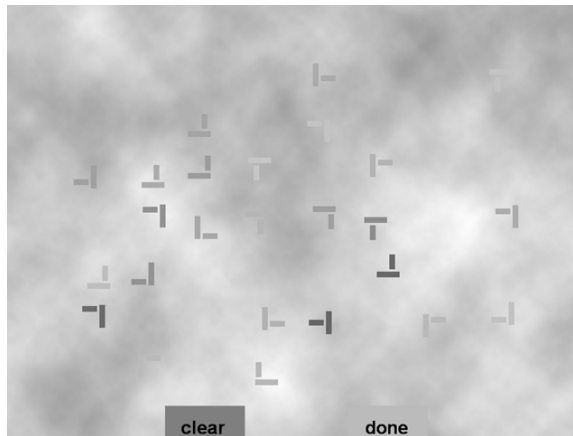


Figure 9: Sample search display for SOS Experiments 1-3, 6-11.

4.1.1.2 Procedure

Each trial began with a cross appearing for 0.5 s at the center of the screen. The cross was replaced with the search array, which consisted of 25 items. Participants were informed that there were either 0, 1, or 2 target Ts to find within each display. Participants used the mouse to click on each detected target item, and then clicked a blue button at the bottom of the screen labeled “DONE” to complete the trial. The “DONE” button appeared 3 seconds after the onset of each trial, and the mouse cursor was reset to the center of the screen after each trial. Participants could correct an error or a mis-click by clicking a yellow button at the bottom of the screen labeled “CLEAR” before completing the trial. Each trial had a time limit of 15 seconds, after which no further clicks were accepted and a message was displayed encouraging participants to try to finish searching and press the “DONE” button before time elapsed on subsequent trials. Responses made prior to the timeout were recorded and analyzed even if the “DONE” button was not pressed.

Trials were classified as one of four types based on the number and the salience of the target Ts presented within the array of distractor Ls, resulting in trial types of: no-target, single high-salience, single low-salience, or dual-target (both a high-salience target and a low-salience target present). Experiments 1-3 only varied in the relative proportion of trial types presented to each set of participants to bias their expectation about how frequently a high-salience or low-salience target might appear. High-salience single targets were as equally frequent as low-salience single targets in Experiment 1, twice as frequent in Experiment 2, and three times as frequent in Experiment 3.

Participants completed 250 test trials with no feedback, divided into 5 blocks of 50 trials each. The experiments had the following distribution of high-salience single-target, low-salience single-target, dual-target, and no-target trial types: Experiment 1: 50, 50, 50, 100; Experiment 2: 100, 50, 50, 50; Experiment 3: 120, 40, 40, 50. Each experiment began with one block of 20 practice trials which were matched to the trial type distribution of the rest of the experiment. Participants were not informed about how often targets would be present, although the practice block gave some indication about what to expect in the rest of the experiment. During the practice block, immediate feedback was provided on any false positive identification or missed targets.

4.1.1.3 Planned analyses

To assess SOS errors, the critical planned comparison for each experiment is how accurately the low-salience targets are detected when they are the only target in the

display (low-salience single targets) versus how accurately the low-salience targets are detected in a dual-target trial, given that the high-salience target was also detected. This calculation of SOS differs slightly from that typically employed in radiology. Radiology studies have often used a small number of trials in which each specific image is presented twice, once with a single target and once with that same target accompanied by an additional target. Such studies typically examine SOS by looking for changes in receiver operating characteristic (ROC) curves, which incorporate a reported confidence rating about each detected target. We do not utilize identical displays with or without an added target and instead compare a large number of single- and dual- target trials which provides the power to look for generalized effects. We also do not include confidence ratings, given the reduced role of decision-making required with our simpler stimuli.

Here we calculate SOS for low-salience targets by conducting planned one-tailed paired-sample t-tests comparing accuracy on single-target low-salience trials with accuracy for the low-salience targets on dual-target trials. We are using 1-tailed tests for this statistical comparison since our a priori SOS prediction, based upon the radiology research, is that detection should be better on single-target than on dual-target trials. Single-target accuracy is measured by the number of hits divided by total number of single-target trials, and dual-target accuracy is measured by the number of dual-target trials in which both targets were detected divided by the total number of dual-target trials in which the high-salience target was detected. Dual-target trials in which both

target types were missed are therefore not included in these analyses. We also conduct a repeated measures ANOVA on these accuracy data to determine effects of frequency, salience, number of targets, and interactions between these factors.

4.1.2 Results and Discussion

We provide detection accuracy here as our primary measure of interest, and we offer response time data in section 4.8. Mean detection rates for high-salience single targets, low-salience single targets, high-salience targets on dual-target trials, and low-salience targets on dual-target trials are presented in Table 1. A repeated measures ANOVA was run with within-subjects factors of Salience (low, high) and Number of Targets (1, 2) and a between-subjects factor of Relative Frequency (1x, 2x, 3x). There was a significant main effect of salience ($F(1,27) = 203.79, p < .001$), a significant main effect of number of targets ($F(1,27) = 5.78, p < .05$), and a significant interaction between number of targets and frequency ($F(2,27) = 4.43, p < .05$). Planned t tests for low-salience targets revealed that the SOS effect (lower accuracy rate for dual-target trials than for single-target trials) was not significant at the 1x frequency (Experiment 1), but was significant at the 2x and 3x frequencies (Experiments 2 and 3; see Table 1).

For all three experiments, the participants had a small proportion of trials in which they did not click "DONE" before the 15 second time limit (Experiment 1: $M=1.4\%$, $SD=1.2\%$; Experiment 2: $M=1.1\%$, $SD=0.6\%$; Experiment 3: $M=1.5\%$, $SD=1.4\%$). As well, there were few false alarms, which were calculated as the percentage of all trials

with one or more false positive responses. (Experiment 1: $M=2.8\%$, $SD=2.4\%$; Experiment 2: $M=1.6\%$, $SD=1.9\%$; Experiment 3: $M=2.3\%$, $SD=3.7\%$).

The results of Experiment 1 indicated that salience differences alone are not enough to trigger SOS errors. That is, in a dual-target situation wherein one target is highly salient and another is less salient and both are equally likely to be present in any given display, detection of the high-salience target does not lead to a higher miss rate for the low-salience target than when the low-salience target is presented by itself.

However, Experiments 2 and 3 increased the frequency of the high-salience targets while keeping constant the frequency of the low-salience targets. The goal was to shift the participants' biases about the type of target to expect on any given trial. Accordingly, this shift in expectation led participants to miss low-salience targets more often on trials in which a high-salience (and highly expected) target was detected than on trials in which the low-salience target was presented alone. Thus, while salience alone was not enough to induce SOS (Experiment 1), it appears that frequency and salience can interact to generate SOS errors. We will return to this result in the General Discussion after exploring additional possible influences on SOS.

Table 1: Experimental parameters and accuracy data for SOS Experiments 1-13.

Expt	Experimental Parameters			Accuracy				SOS Effect	
	Stimuli	Time Limit	Prevalence (High:Low)	High single	High dual	Low single	Low dual	Low-single vs. Low-dual	SOS ?
1	Ts & Ls	15s	1:1	88.40% (8.42)	90.19% (7.63)	60.40% (12.57)	63.09% (10.72)	$t(9)=1.28$ $p=0.233$	✗
2	Ts & Ls	15s	2:1	94.80% (5.57)	90.02% (9.72)	66.00% (16.22)	59.72% (16.62)	$t(9)=3.18$ $p=0.011$	✓
3	Ts & Ls	15s	3:1	95.66% (3.42)	96.07% (3.83)	72.50% (12.42)	60.76% (12.23)	$t(9)=3.56$ $p=0.006$	✓
4	Pictures	15s	1:1	90.00% (5.59)	89.70% (11.55)	72.02% (13.30)	75.37% (19.45)	$t(9)=1.46$ $p=0.180$	✗
5	Pictures	15s	9:1	82.00% (8.82)	82.86% (12.74)	59.39% (12.38)	57.18% (20.04)	$t(9)=0.96$ $p=0.360$	✗
6	Ts & Ls	30s	2:1	94.20% (4.10)	92.64% (6.55)	60.60% (17.26)	58.50% (19.59)	$t(9)=1.07$ $p=0.313$	✗
7	Ts & Ls	30s	3:1	95.74% (2.35)	92.13% (8.25)	60.50% (13.83)	55.64% (16.67)	$t(9)=2.08$ $p=0.067$	~✓
8	Ts & Ls	30s	6:1	93.26% (5.77)	96.08% (5.48)	68.00% (12.51)	57.90% (15.98)	$t(9)=2.11$ $p=0.064$	~✓
9	Ts & Ls	30s	1:1 (Low-single:dual)	-	-	78.17% (7.80)	76.75% (13.22)	$t(9)=1.12$ $p=0.292$	✗
10	Ts & Ls	30s	4:1 (Low-single:dual)	-	-	74.91% (8.70)	71.63% (5.43)	$t(9)=3.10$ $p=0.013$	✓
11	Ts & Ls	30s (Instructions)	3:1	94.26% (4.55)	95.64% (4.07)	62.75% (15.48)	62.04% (17.55)	$t(9)=0.86$ $p=0.414$	✗
12	Ts & Ls	Luggage Line (variable rate)	3:1	97.10% (4.60)	95.32% (6.12)	74.07% (15.11)	73.08% (11.84)	$t(9)=0.93$ $p=0.379$	✗
13	Ts & Ls	Luggage Line (constant rate)	3:1	84.91% (7.81)	76.22% (16.15)	71.89% (12.77)	63.92% (15.11)	$t(9)=3.21$ $p=0.011$	✓

4.2 Experiments 4-5: Examining effects of decision-making processes on SOS

Radiography studies have delineated three possible types of errors in which SOS may occur: scanning errors (the search path never encounters the target area), recognition errors (scanning in the region of a possible target but failing to dwell on the correct area for further inspection), and decision-making errors (fixating and dwelling on a possible target but ultimately failing to identify it as a target). To date, evidence has suggested that all three possibilities, scanning errors (Berbaum et al., 1996; Samuel et al., 1995), recognition errors (Berbaum et al., 2000), and decision errors (Franken et al., 1994), contribute. The latter two explanations differ primarily in the amount of time spent analyzing a potential target. The need to spend extra time to examine a target stems in part from the relatively low spatial frequency of radiological targets (e.g., abnormalities, pulmonary nodules) which may require extra analysis to visually parse them from background noise. By categorizing errors as scanning, recognition, or decision errors, radiologists have attempted to understand whether SOS arises primarily as a function of a basic perceptual failure to properly scan an image, a failure of pattern recognition, or a failure of knowledge-based decisions.

To add to this discussion, in Experiments 4 and 5 we explored the role of analysis time on SOS. We changed the nature of the stimuli to reduce the decision-making component of visual search and to enable rapid recognition of targets. Whereas the Ts and Ls of Experiments 1-3 demand extra analysis in many cases since the distractor Ls

were often minimally distinguishable from target Ts, in Experiments 4 and 5 we utilized a small set of photographic images to facilitate rapid recognition and to minimize the per-item decision-making process. In addition, the photographic stimuli better match the high spatial frequency of targets which might be found in an airport security luggage screening task.

4.2.1 Methods

Except where noted the methods were identical to those of Experiments 1-3.

4.2.1.1 Participants and stimuli

Twenty individuals (Experiment 4: mean age=19.5 years, $SD=1.2$ years; Experiment 5: mean age=18.6 years, $SD=0.7$ years) from the Duke University community participated (10 in each experiment). All participants gave informed consent and received either course credit or \$10/hour.

Stimuli consisted of 30 items drawn from a pool of 240 photographs of *sports equipment, toys, foods, clothing, birds, musical instruments, tools, and bottles* (average item subtended $3.8^\circ \times 3.8^\circ$ total). The photographs were drawn from the Hemera Photo-Objects Collections (Hemera Photo Objects, Gatineau, Quebec, Canada) and have previously been used in similar search tasks (e.g., Fleck & Mitroff, 2007; Wolfe et al., 2005). Each object was converted to grayscale and partially blurred, then presented with a random rotation within randomly selected cells of an invisible 6×5 grid subtending $20.3^\circ \times 15.2^\circ$ at an approximate viewing distance of 60 cm. Items were close enough to

partially overlap with surrounding items, and all were partially transparent. Stimuli were presented on a rendered grayscale “cloud” background (brightness range=10-50 % black) that differed on each trial (Figure 10). Distractor items (*toys, foods, clothing, birds, sports equipment, musical instruments*) were presented at a range of transparencies (range=20-60 % black). Target photos were items belonging to either the *bottle* category or the *tool* category and were presented either at high-salience (range=45-65 % black) or low-salience (range=20-40 % black). Except where otherwise noted, *bottles* were presented at high-salience, and *tools* were presented at low salience.

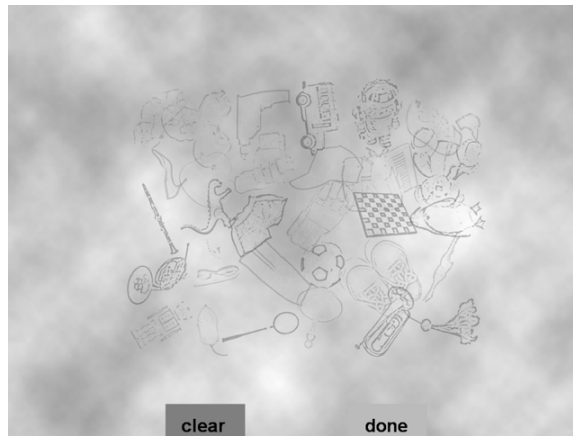


Figure 10: Sample search display for SOS Experiments 4 and 5.

4.2.1.2 Procedure

The goal of Experiments 4 and 5 was to determine how stimulus type would modulate the SOS effect and whether the previously noted salience and frequency interplay would again predict multiple target errors. Therefore, the search procedure was nearly identical to Experiments 1-3. In Experiment 4, the frequency of high-salience

targets equaled the frequency of low-salience targets, whereas in Experiment 5 the high-salience targets were nine times more frequent than low-salience targets.

Given the nature of the stimuli and the large difference in high-salience and low-salience target frequency for Experiment 5, several other minor procedural differences were introduced. In Experiment 4, participants completed 50 practice trials with feedback, then 270 test trials with 45 single high-salience trials, 45 single low-salience trials, 30 dual-target trials consisting of two high-salience targets, 30 dual-target trials consisting of one high-salience target and one low-salience target, 30 dual target trials consisting of two low-salience targets, and 90 no-target trials. In Experiment 5, participants completed 50 practice trials with feedback, then 600 test trials with 360 single high-salience trials, 20 single low-salience trials, 20 dual-target trials consisting of one high-salience target and one low-salience target of the same category (both *tools* or both *bottles*), 20 dual-target trials consisting of one high-salience target and one low-salience target of different categories (one *tool* and one *bottle*), and 180 no-target trials. Experiment 5 consisted of 10 blocks, divided into two 1-hour sessions conducted on consecutive days. Several trials types were employed to equate and manipulate various distributions, but here we restricted our SOS analyses to the same trial types examined in Experiments 1-3: single low-salience trials and dual-target trials (both a high-salience target and a low-salience target present).

4.2.2 Results and Discussion

Mean detection rates are presented in Table 1 (see Table 2 in section 4.8 for response time data). A repeated measures ANOVA was conducted with within-subjects factors of Saliency (low, high) and Number of Targets (1, 2) and a between-subjects factor of Relative Frequency (1x, 9x). There was a significant effect of saliency ($F(1,18) = 83.32, p < .001$), but no significant effect of number of targets ($F(1,18) < 1, p > .05$), and no significant interaction between number of targets and frequency ($F(1,18) < 1, p > .05$). Planned *t*-tests for SOS at low-saliency showed no differences for Experiment 4 or 5 (see Table 1). Despite the change in stimuli, Experiment 4 replicated Experiment 1, producing no SOS effect when the frequency of low-saliency targets matched the frequency of high-saliency targets. However, unlike Experiments 2 & 3, increasing the frequency of high-saliency targets relative to low-saliency targets in Experiment 5 did not accordingly modulate the SOS effect.

There were a minimal number of trials in which participants did not click “DONE” before the 15 second time limit (Experiment 4: $M=2.2\%$, $SD=2.1\%$; Experiment 5: $M=1.3\%$, $SD=1.6\%$). There were few false alarms (Experiment 4: $M=2.0\%$, $SD=1.9\%$; Experiment 5: $M=5.5\%$, $SD=2.7\%$).

When participants searched for recognizable and easily distinguished photographic targets, the detection of one target did not interfere with the detection of an additional target, and this null result held even when the occurrence of the low-

salience target was extremely rare relative to the high-salience target. These data implicate that a decision-making component may be necessary for SOS, even when considering the frequency and salience factors shown to generate SOS in Experiments 1-3. Although no SOS was found here, caution needs to be taken when generalizing this null result to real-world searches: These search arrays had 30 total items and targets that repeated throughout the experiment, whereas many real-world searches have a larger set size and an undefined set of possible targets. A larger search space might lead to the type of SOS errors found in radiological studies which have evidenced incomplete scanning (e.g., Samuel et al., 1995) and faulty recognition (e.g., Berbaum et al., 2000).

The photographic stimuli used in these experiments introduced an additional, potentially important difference in the visual search task: a multiple category factor. In Experiments 1-3 participants searched for Ts which could be either high-salience or low-salience, but except for salience the actual T target shapes were identical. In Experiments 4 and 5, high-salience targets were primarily of one category (*bottles*) and low-salience targets were primarily of another (*tools*), and in all cases, there were several unique exemplars. Not only did participants have to search for two category types, but they could no longer look for one specific pattern to match. Interestingly, in spite of this added load of searching across different semantic categories, participants showed no adverse effect for detecting less frequent targets when finding a more frequent and salient target. However, it should be noted that this paradigm utilized a finite set of

potential targets (30 possible tools and 30 possible bottles) to which participants had some exposure in the practice block. It will be interesting for future work to explore whether SOS is sensitive to these issues of target heterogeneity (the number of target categories and diversity of patterns within those category) and familiarity (whether or not the participant has seen in training or practiced with a perceptually similar prototype of that target) when the pool of possible targets is much more diverse, such as in an actual airport security screening task.

4.3 Experiments 6-8: Examining the effect of time pressure on SOS

Radiological studies of SOS are typically self-paced paradigms modeled after routine radiograph examinations (e.g., Berbaum et al, 1998). Even while such studies have demonstrated significant SOS effects, no study has yet to directly test the effects of a specific time pressure which might be presented in real-world radiographic workflows. A radiologist may understand the need to process a minimum number of radiographs in a day, a pressure which is absent in a controlled radiology experiment. Similarly, during high passenger flow, luggage screening typically occurs on the order of 3-5 seconds per inspection (Schwaninger, 2005), indicating the possibility of a stricter time pressure than that found in radiology. Here we ask whether time pressure plays a role in the SOS effect.

In Experiments 1-5, we utilized a 15 second time limit per trial. Although real-world search tasks may indeed be faced with time pressures, we wish to establish

whether the SOS effects observed in Experiments 2 and 3 were potentially driven by time pressure, rather than simply an interplay between salience and frequency. In Experiments 6-8 we double the amount of time allowed to search on each trial. We again manipulate the frequency of salient targets to explore the relationship between salience, frequency, and time pressure.

4.3.1 Methods

Except where noted the methods were identical to those of Experiments 1-3. Thirty individuals (Experiment 6: mean age=18.6 years, *SD*=1.0 years; Experiment 7: mean age=19.1 years, *SD*=1.2 years; Experiment 8: mean age=18.8 years, *SD*=1.4 years) from the Duke University community participated (10 in each experiment). In Experiments 6-8, each participant was allowed 30 seconds to search each display, which was twice as long as allowed in Experiments 1-5. The difference between Experiments 6, 7, and 8 was the frequency with which high-salience targets would appear: In Experiment 6, high-salience targets were 2x more frequent than low-salience targets; in Experiment 7, high-salience targets were 3x more frequent; and in Experiment 8, high-salience targets were 6x more frequent. The experiments had the following high-salience single-target, low-salience single-target, dual-target, and no-target trial type distributions: Experiment 6: 100, 50, 50, 50; Experiment 7: 120, 40, 40, 50; Experiment 8: 150, 25, 25, 50.

4.3.2 Results and Discussion

Mean detection rates are presented in Table 1 (see Table 2 in section 4.8 for response time data). A repeated measures ANOVA was conducted with within-subjects factors of Saliency (low, high) and Number of Targets (1, 2) and a between-subjects factor of Relative Frequency (2x, 3x, 6x). There was a significant effect of saliency ($F(1,27) = 170.01, p < .001$), a significant effect of number of targets ($F(1,27) = 4.49, p < .05$), and no significant interaction between number of targets and frequency ($F(2,27) < 1, p > .05$). Planned follow-up *t*-tests for low-saliency targets revealed no SOS effect for Experiment 6 and weak effects for Experiments 7 and 8 (see Table 1). The doubled available time sharply reduced the SOS effect but did not eliminate it altogether. Although Experiments 6 and 7 were identical to Experiments 2 and 3, respectively, in every other respect except for time limit, there was no sign of SOS in Experiment 6 and only a weak effect in Experiment 7. Even when high-saliency targets were 6x more likely than low-saliency targets in Experiment 8, there was only weak evidence for the SOS effect.

There was a vanishing proportion of trials in which participants did not click “DONE” before the 30 second time limit (Experiment 6: $M=0.2\%$, $SD=0.3\%$; Experiment 7: $M=0.2\%$, $SD=0.3\%$; Experiment 8: $M=0.1\%$, $SD=0.3\%$). As well, there were few false alarms (Experiment 6: $M=1.4\%$, $SD=1.5\%$; Experiment 7: $M=1.6\%$, $SD=2.3\%$; Experiment 8: $M=3.1\%$, $SD=3.7\%$).

The effect of doubling the available search time for each trial had a significant impact, reducing SOS errors in all three experiments. Importantly, reducing the time pressure here eliminated SOS for an otherwise identical condition that had previously yielded SOS (Experiment 2 vs. Experiment 6). However, SOS was not entirely eliminated by the additional time in Experiments 7 and 8; the factors of salience and frequency continued to play a role in hindering multiple target detection, even with the reduced time pressure. These results indicate that real-world tasks with potentially more than one target would strongly benefit from striving to reduce the time pressure on the searcher, whether the task is in radiology, cytology, or airport security. Extra time may not eliminate SOS entirely, but time pressure clearly serves to exacerbate the problem of accurately detecting multiple targets.

4.4 Experiments 9-10: Examining how identical salience in additional targets modulates SOS

Experiments 1-8 explored the influences of salience, frequency, and time pressure on multiple target accuracy, and they revealed a key role of decision-making processing in producing SOS. To examine these issues, each of the experiments so far has utilized two types of targets: high-salience and low-salience. This was motivated by salience-related SOS effects in medical research (Berbaum et al., 1994; DeMay, 1997) and an attempt to map onto the saliency/frequency differences found in current airport security screening. However, we also wished to determine the extent to which SOS may arise when a second target is identical to the first detected target.

It has been suggested that SOS may arise in part as a consequence of radiologists adopting a readiness to seek out a specific pattern in the image (Berbaum et al., 1990). In such a “perceptual set” explanation, the detection of one target type, say, a tumor of a particular contrast, may bias searchers to selectively search for additional targets with the same perceptual features and to discount the features of targets from different categories (Berbaum et al., 1990). We look to explore this possibility here given our observed effects of expectations on SOS in Experiments 1-8 and given the robust “attentional set” literature which confirms such predictions in basic cognitive psychology experiments (e.g., Folk, Remington, & Johnston, 1992; Most, Scholl, Clifford, & Simons, 2005). If SOS is influenced by something akin to a “perceptual set” such that expectations cause participants to bias the search for one stimulus type (e.g., high-salience) over the other (e.g., low salience), then the effect could be eliminated when both targets are identical. Additionally, in most real-world tasks, additional targets are often identical to the primary target, and so we wish to establish if SOS can arise in this general situation.

In Experiment 9, we remove all expectation biases about both the number of targets and the target salience by making secondary targets identical to the primary target and by making dual-target trials as equally likely as single-target trials. Thus, on any given trial, it is equally likely that there may be one or two targets in the display, and if there are two targets, they are identical in salience and shape. In Experiment 10,

we keep the identical-target manipulation while reintroducing the element of frequency and expectation by altering the ratio of single-target to dual-target trials. Specifically, single-target trials are four times more likely than dual-target trials. The goal is to determine if SOS is sensitive to expectations about the number of targets likely to be present even when all targets are identical.

4.4.1 Methods

Except where noted the methods were identical to those of Experiments 6-8. Twenty individuals (Experiment 9: mean age=19.9 years, $SD=1.0$ years; Experiment 10: mean age=21.3 years, $SD=5.0$ years) from the Duke University community participated (10 in each experiment). To minimize the contribution of the time pressure effect to SOS evidenced by the first 8 experiments, we continued to use a 30 second time limit.

All targets were presented at the low-salience parameters of Experiments 1-3, resulting in three trial types: no-target, single low-salience, and dual low-salience. The experiments had the following distribution of low-salience single-target, low-salience dual-target, and no-target trial types: Experiment 9: 100, 100, 50; Experiment 10: 160, 40, 50.

4.4.2 Results and Discussion

Mean detection rates for low-salience single-target trials and low-salience dual-target trial are presented in Table 1 (see Table 2 in section 4.8 for response time data). We conducted a repeated measures ANOVA with a within-subjects factor of Number of

Targets (1, 2) and a between-subjects factor of Single-to-Dual Frequency (1x, 4x). There was a trend towards significance for the main effect of number of targets ($F(1,19) = 2.86$, $p = .11$), and no significant interaction between number of targets and trial type frequency ($F(1,19) < 1$, $p > .05$). Planned follow-up t tests for low-salience targets revealed no SOS effect for Experiment 9, as expected, but in Experiment 10 when single-target trials were four times more likely to occur than dual-target trials, participants demonstrated a significant SOS effect (see Table 1). There was a small proportion of trials in which participants did not click “DONE” before the 30 second time limit (Experiment 9: $M=0.5\%$, $SD=0.9\%$; Experiment 10: $M=0.2\%$, $SD=0.3\%$). As well, there were few false alarms (Experiment 9: $M=5.6\%$, $SD=5.9\%$; Experiment 10: $M=1.5\%$, $SD=1.1\%$).

By removing the factor of differing target salience levels, we observed the expected result that SOS was abolished. This finding from Experiment 9 again emphasizes that differing target salience, and expectations about those differing levels, is an important factor in predicting SOS. Moreover, when we increased the frequency of single-target trials relative to dual-target trials, thus creating an expectation of most trials containing only one target, an SOS effect emerged wherein participants missed the secondary target more often when paired with another identical target than when presented alone. These experiments indicate that a “perceptual set” account cannot

entirely drive SOS since even when all targets are perceptually alike, detection of a secondary target is diminished simply by the act of successfully finding the first target.

4.5 Experiment 11: Examining the effect of instruction on SOS

Is it possible that specified search strategies and/or a priori top-down knowledge can modulate SOS? In radiology, the use of “checklists” has been recommended as an external aid to offer protection against SOS errors during the radiographic search (Kinard, Orrison, & Brogdon, 1986; Samuel et al., 1995), although more recent evidence suggests that this may actually increase the likelihood of SOS (e.g., Berbaum, Franken, Caldwell, & Schartz, 2006). In addition, other work has shown that incorporating knowledge about the clinical history of a particular case can offset SOS errors. Here we explore this issue of external guidance by offering participants information about the task they are about to perform.

Experiments 1-8 collectively suggest that SOS is affected by participants' expectation about the relative frequency of high-salience and low-salience targets and Experiments 9 and 10 likewise suggest that SOS is affected by an expectation about the frequency of single- and dual-target trials. We predict here that explicitly drawing the participants' attention to the relative frequency of high-salience and low-salience targets might lead to a strategic avoidance of the SOS effect. In this experiment, we informed participants about the increased likelihood of high-salience targets and instructed them to find those targets first before searching for additional targets. We predicted that the

SOS effect may be attenuated by an explicit highlighting of the parameters of the search which generate SOS.

4.5.1 Methods

Ten individuals (mean age=18.4 years, $SD=1.2$ years) participated in the study and it was identical to Experiment 7 in all regards except for a change in instructions prior to the task. Participants were verbally informed that “Dark Ts are much more likely to occur than light Ts, and thus a good strategy for the task is to first search for any easy-to-see dark Ts, then search for the more difficult targets which might also be present.”

4.5.2 Results and Discussion

As seen in Table 1, the verbal instructions eliminated any evidence of the SOS effect. In fact, the data suggest a reversed effect where participants were slightly better able to detect the low-salience targets when a high-salience target was found than when the low-salience target was presented alone. A repeated measures ANOVA was conducted with the factors of Salience (low, high) and Number of Targets (1, 2). There was a significant effect of salience ($F(1,9) = 39.92, p < .001$) but no significant effect of number of targets ($F(1,9) < 1, p > .05$). There were no trials in which participants did not click “DONE” before the 30 second time limit and there were few false alarms ($M=1.4\%$, $SD=1.4\%$). See section 4.8 for response time data.

The current consensus in radiology is that a checklist may lead to increased SOS (Berbaum et al., 2006). Yet, strategic knowledge about what to expect while searching (i.e., clinical history of the case) has also been shown to reduce SOS. Here, we found that specific instructions about how to perform a visual search, based on actual probabilities of the targets, reduced the SOS effect. Interestingly, this suggests that although expectation about target types can induce SOS, metacognition about this expectation (i.e., the degree to which a searcher is explicitly aware of what to expect) may serve to boost awareness of less frequent target types. This role of metacognition also poses a potentially interesting wrinkle while examining frequency effects in paradigms without instruction: Tasks with less-noticeable frequency disparities (e.g., 2x or 3x) may lead to less metacognition about such differences and thus actually predict greater SOS than in paradigms with larger frequency differences (e.g., 9x), which may highlight during practice trials (or real-world training) the need to pay particular attention to rarer target types.

4.6 Experiments 12-13: Examining the interaction of time and reward pressures on SOS

Experiments 1-11 have collectively demonstrated several constraints that influence SOS, including an effect of time pressure. However, these laboratory-based experiments with novice searchers obviously do not entirely reproduce the host of other pressures that a real-world searcher faces. For example, a radiologist, a cytologist, and an airport X-ray screener all have an immeasurably higher incentive to avoid missed

targets than participants volunteering in a laboratory study. Although it is extremely difficult to replicate in the lab such real-world pressure to perform accurately, in Experiments 12 and 13, our goal was to boost the motivation of participants by rewarding successful target detection to determine if SOS is sensitive to the effort and motivation devoted to the search task. Even while radiologists, who are certainly highly motivated, demonstrate SOS, we wish to establish that the generalized SOS effects found in Experiments 1-11 are not simply driven by a low motivation to find all targets. In many laboratory studies there is a concern that participants may wish to complete the experiment as quickly as possible, and in a multiple target paradigm this is a particularly hazardous prospect. Participants may feel they have “adequately participated” on each trial upon finding any target, thereby manifesting the exact property we wish to explore. Although arguably this prospect is inherent as well in any non-laboratory search task, here we ask whether the SOS effect may be attenuated when motivation to find all possible targets is increased.

Moreover, we wanted to allow participants the ability to “budget” their time across the length of the experiment to enable them to selectively balance accuracy versus time. In many real-world tasks, searchers are not subject to the “fixed” time limit pressure as we have utilized so far. Instead, they have the option of economically allocating time appropriately—spending more time on the cases which require additional attention while speeding through images deemed “easier” upon initial

inspection. In this manner, the time pressure is less a per-image pressure as it is a per-session pressure (e.g., to search a given number of images within a day). In Experiments 12 and 13, we implement a performance-based reward system simulating a luggage screening task to see how and when SOS would arise when participants are motivated to perform accurately, and are facing a time pressure across the experiment rather than within a trial.

4.6.1 Methods

Except where noted the methods were identical to those of Experiment 3. Twenty individuals (Experiment 12: mean age=18.7 years, $SD=0.9$ years; Experiment 13: mean age=19.1 years, $SD=1.1$ years) from the Duke University community participated (10 in each experiment). Participants were paid \$10 with the possibility of an extra \$10 for high performance (see below).

The main search task was identical to Experiment 3, except we implemented an accuracy- and time-based reward system for performance on the task rather than a fixed time limit. Participants were instructed to manage a “line of luggage,” represented by a row of luggage icons at the top of the screen above the searched area, which would increase and decrease in the number of icons throughout the experiment (see Figure 11). As participants completed each trial, regardless of accuracy on the trial, one icon would be eliminated from the row. Icons would be added to the row as a function of time spent on each trial. In Experiment 12, this time window was set individually for each

participant based on their average time spent on no-target trials in the practice block. Slower participants would therefore accrue “luggage icons” at a slower rate than faster participants, and this manipulation reduced the time pressure factor overall while still demanding that participants budget accuracy versus time spent searching. In Experiment 13, this time window for luggage accrual was fixed for all participants at 9 second per icon to establish a much greater time pressure and to determine whether participants would adapt to the relatively fast rate of luggage accrual.

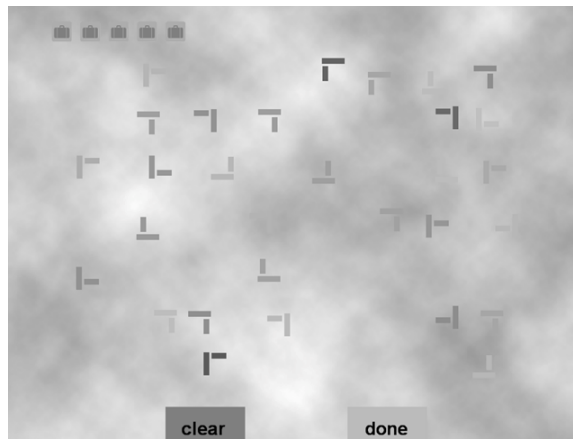


Figure 11: Sample search display for SOS Experiments 12 and 13.

The luggage line was tied to a point-based reward system. Participants were told that successfully responding to any trial (either with zero, one, or two targets) would always result in a gain of points, but that these points would be greater if the luggage line at the top of the screen was kept short. Participants were warned that missed targets would result in a significant loss of points, and that even when the luggage line reached its longest length (16 icons along the top of the screen), participants would continue to

gain points for each correct trial. Participants were told that accuracy was the primary goal in the task, and that line-length management was a secondary goal. The computer tracked each participant's cumulative point total offscreen, and participants were informed that the top points-scorer out of every three participants would receive an extra \$10. After each block, participants were informed about their current point total, but they were never told about the total points of any previous participant.

Participants completed 50 practice trials with feedback, followed by 200 test trials with no feedback, divided into 4 blocks of 50 trials each. Each display contained 30 items. Similar to Experiments 3 and 7, we utilized a high-salience:low-salience single target trial ratio of 3:1 to replicate known SOS conditions. Both Experiments 12 and 13 had 96 high-salience single-target trials, 32 low-salience single-target trials, 32 dual-target trials, and 40 no-target trials.

4.6.2 Results and Discussion

Mean detection rates are presented in Table 1 (see section 4.8 for response time data). A repeated measures ANOVA was conducted with within-subjects factors of Salience (low, high) and Number of Targets (1, 2) and a between-subjects factor of Time Pressure (variable, fixed). There was a significant effect of salience ($F(1,18) = 73.28, p < .001$), a significant effect of number of targets ($F(1,18) = 11.27, p < .005$), a significant effect of time pressure ($F(1,18) = 5.60, p < .05$), a significant interaction between salience and time pressure ($F(1,18) = 5.85, p < .05$), and a significant interaction between number of

targets and time pressure ($F(1,18) = 5.76, p < .05$). There was a small proportion of false alarms in each experiment (Experiment 12: $M=0.7\%$, $SD=1.0\%$; Experiment 13: $M= 2.1$, $SD= 1.8\%$). Experiment 12, with the rate of luggage accrual customized for each participant as a function of their practice block speeds, mirrored the general accuracy rate of Experiment 3, but did not exhibit an SOS effect. Alternatively, Experiment 13, with the rate of luggage accrual fixed at 9 seconds across all participants, produced lower accuracies overall and a significant SOS effect (see Table 1 for data and planned t -test).

When compared with Experiments 3 and 7, which yielded evidence of SOS, the lack of a significant SOS effect in Experiment 12 may indicate that motivation indeed plays a role. However, it is difficult to interpret the results directly, as the addition of the luggage-management task as a secondary goal may have changed the primary task too significantly to directly compare. This is particularly important to acknowledge since we did observe an SOS effect in Experiment 13, suggesting that even when motivation is high (here, in the form of possible financial reward), the pressures of the task including salience, frequency of target types, and time pressure can still interact to adversely affect the ability to detect additional targets.

4.7 General Discussion

The 13 experiments presented here focused on the phenomenon of “satisfaction of search” (SOS), whereby the detection of one target is hindered by the successful

detection of another target. This research was conducted with the goal of establishing the commonalities between visual search as it is studied in radiology and how it is studied in cognitive psychology. Since SOS may not be a problem exclusive to radiology, we have investigated here the generality of the effect and explored the parameters to which it is sensitive. By manipulating a variety of search parameters we have revealed several contexts in which SOS errors can be observed, even in a non-medical search task, and with novice searchers. Importantly, this general finding of SOS outside of radiology suggests that the factors which modulate accuracy in multiple target detection may be broadly applicable to many other critical, real-world search tasks such as airport security screening. These factors include: (1) the relative salience and frequency of different target types, (2) stimulus discriminability, (3) time pressure, (4) perceptual set, (5) search instructions, and (6) reward pressure.

4.7.1 Summary

Experiments 1-3 established that SOS is sensitive to the interaction between salience and the expectancy about certain target events: As the frequency of easy-to-detect high-salience targets increased relative to difficult-to-detect low-salience targets, participants missed more of those low-salience targets in a dual-target condition than when the low-salience target was presented by itself. Interestingly, salience differences alone (Experiment 1) were not sufficient to induce the SOS effect. Rather, it is the interaction between those salience differences and biased expectation about the differing

target types which leads to SOS; in Experiments 2 and 3, we shifted the frequency of targets such that participants expected high-salience targets more often than low-salience targets and consequently dual-target low-salience performance fell.

Experiments 4 and 5 were conducted to determine if a prolonged decision-making process is necessary for inducing SOS, and our data suggest that this may indeed be the case. When we switched our stimuli from Ts and Ls, which require a careful serial search and analysis, to a set of photographs, which can be more rapidly identified once fixated, we failed to observe SOS (Experiment 4). Even when the factor of frequency was greatly increased, a parameter shown to induce SOS in the first three experiments, the photographic stimuli still failed to induce SOS in participants (Experiment 5).

In Experiments 6-8 we looked to establish the role of time pressure in SOS. A potential concern with the significant SOS effects in Experiment 2 and 3 is that a 15 second time limit may not be enough time to adequately search the display and find two targets. In Experiments 6-8 we doubled the time limit to 30 seconds, and although this weakened the SOS effect, it was not eliminated, thus revealing time pressure as another contributing factor to SOS.

In Experiments 9 and 10 we focused more closely on the necessity of salience differences between multiple targets. We wished to determine if SOS can arise even when all targets are perceptually identical. A failure to find an effect when targets are

perceptually similar would suggest that SOS arises as a function of a “perceptual set” which biases the decision to terminate a search after the “preferred target” has been detected. However, even when targets were perceptually identical we still observed SOS, although this effect still relied on expectancy bias: SOS only emerged when dual-target trials were less frequent than single-target trials.

In Experiment 11, we showed that instructions to explicitly attend to a particular interaction which causes SOS (salience and frequency) can attenuate the effect. Lastly, in Experiments 12 and 13 we implemented a performance-based reward system to determine if increased motivation can offset the SOS effect. That is, if participants are given an incentive to be both accurate and fast, will they no longer reveal SOS? Whereas Experiments 3 and 7 demonstrated a SOS effect, Experiment 12 (which had comparable parameters) did not. This reduced SOS effect could have resulted from the increase in motivation, but future work may be needed to determine the added influences of managing an onscreen luggage queue while searching for targets. Importantly, when time pressure was increased in Experiment 13, participants demonstrated a strong SOS effect, suggesting that increased motivation alone is insufficient to counter the tendency to miss additional targets in visual search.

4.7.2 Implications for SOS in radiology

SOS has been extensively studied in radiology and the present experiments attempted to establish the generality of the effect to determine if and how SOS errors

arise in non-medical domains. However, these generalized results may in turn inform the attempts to minimize SOS within radiology. First, the present data add to the debate about the contribution of scanning errors, recognition errors, and decision errors to SOS effects by indicating that SOS is particularly sensitive to the discriminability of a target and the process of target recognition. When participants were able to more swiftly recognize photographic targets rather than study the details of a possible T shape, SOS was abolished, and this effect held even when participant bias was heavily influenced by a drastic manipulation of relative target frequency. While our photographic images and T shapes are both far from the complexity of radiographic images, their relative effect on SOS here is potentially quite informative.

Second, although target frequency did not induce SOS with photographic stimuli, the finding that SOS is clearly linked to salience and frequency with less discriminable targets (Ts and Ls) may have implications for radiology. Specifically, the theory that a “perceptual set” may act to increase SOS errors is supported by our data, which illustrated that when participants expected a particular visual pattern for the target (in the present studies, a high-salience T shape), the result was that successfully finding such a target interfered with detection of a perceptually different target (here, a low-salience T shape). Similarly, if radiologists adopt a particular readiness to interpret images in a certain context (e.g., a specific kind of abnormality, or a particular shape or contrast), based on recent findings or other expectations, this bias may lead to increases

in SOS errors. Interestingly, although semantic category and perceptual appearance vastly differed between the two target types in Experiments 4 and 5, even a heavily-based expectation for one target type did not induce the SOS effect, emphasizing again the possible role of target discriminability in SOS.

Third, although radiologists may be well trained in efficiently optimizing the trade-off between accuracy and the need to process at least a certain number of cases in a particular session, our data linking time pressure with SOS, both within a particular case (e.g., Experiment 2 vs. 6) and across a longer session (Experiment 12 vs. 13), should be taken into consideration. Although it may be an obvious conclusion, the interaction between expectation and time constraints specifically emphasizes that radiological searches should minimize pressure to process a particular number of cases in a day.

4.7.3 Implications for cognitive psychology

Visual search is a well-studied cognitive task and this is not the first comparison of searches for one versus multiple targets (e.g., Gibson, Li, Skow, Brown, & Cooke, 2000; Körner & Gilchrist, 2008; Menneer et al., 2007; Metlay, Sokoloff, & Kaplan, 1970; Takeda, 2004). However, the influence of multiple possible targets in cognitive studies of search has primarily explored the impact on search efficiency, or the added cost of extra targets on time-to-search, rather than search accuracy, as we examined here. For example, Menneer et al. (2007) nicely showed that the efficiency costs associated with simultaneously searching for two different targets was greater than the sum of the costs

associated with searching for those targets separately. Here, we manipulate the number of targets simultaneously present on any given trial and focus primarily on the accuracy effects. The search parameters we explored in this paper, including salience, relative frequency of events, time pressure, reward pressure, stimulus type, and perceptual set, are not novel in visual search studies, but the current experiments offer some of the first evidence about how these factors interact in a multiple target paradigm to modulate accuracy.

More generally, the finding of SOS in a non-medical context and in an untrained, naïve group of participants suggests that this type of error may reflect a global search heuristic. Indeed, it may be adaptive to exhibit SOS as a means to maximize efficiency of time, just as the phenomenon of “inhibition of return” (Posner & Cohen, 1984) is thought to maximize the efficiency of search and foraging behavior by deprioritizing previously-attended locations. Regardless of the mechanism, that SOS is a general effect, sensitive to pressures non-specific to radiology, suggests that any visual search incorporating multiple simultaneously present targets must take into account the interactions we have demonstrated here, adding to the previously known data on visual search including effects of item frequency (e.g., Fleck & Mitroff, 2007; Wolfe et al., 2005; Rich et al., 2008, Van Wert et al., 2009), familiarity (e.g., Wang et al., 1994), set size (e.g., Carter, 1982; Wolfe, 1998), target and distractor heterogeneity (e.g., Nagy & Thomas, 2003; Palmer, Verghese, & Pavel, 2000), emotionality of stimuli (e.g., Gerritsen, Frischen, Blake, Smilek,

& Eastwood, 2008), and memory (e.g., Horowitz & Wolfe, 1998; Körner & Gilchrist, 2008), to name a few.

4.7.4 Implications for luggage screening and other real-world searches

Given the generality of SOS, these results motivate a careful analysis of real-world search tasks to determine if similar multiple target effects may be contributing to overall error rates in such socially critical searches. In the particular case of airport security X-ray screening, it is obviously a vital safety issue to insure maximal detection rates in luggage screening, yet there has been relatively little crosstalk between radiology, cognitive psychology, and transportation security, despite the many factors common to visual search tasks in all three domains (for some examples, see Fiore, Scielzo, & Jentsch, 2004; Fleck & Mitroff, 2007; Gale et al., 2000; McCarley et al., 2004; McCarley & Carruth, 2004; Menneer et al., 2007; Smith, Redford, Washburn, & Tagliatela, 2005; Wiegmann, McCarley, Kramer, & Wickens, 2006; Wolfe et al., 2005).

Most of the factors explored in this paper can directly inform real-world search tasks such as airport security. Notably, the interaction between salience and expectancy is directly relevant to current security searches. As of this writing, current security protocols mandate that airport screeners search for very salient and common targets such as water bottles, hair gel, soft drinks, and toothpaste, potentially at the expense of finding additional targets which may be better concealed and less frequent, such as scissors, box cutters, or pocketknives. Although extensive training assuredly emphasizes

an exhaustive search no matter the search results, the finding of salience effects in osteoradiology (Berbaum et al., 1994) and cytology (DeMay, 1997) indicates that training may not override this basic search heuristic.

We also found that time pressure interacts with salience and frequency to exacerbate the SOS effect. Critically, airport security searches are much shorter and more numerous in a session than radiograph examinations. Although searchers have no “fixed time limit” after which they can no longer search, there is likely to be a global pressure to keep security lines moving efficiently, thus making these effects of time pressure particularly relevant. Furthermore, although we found some evidence that reward may attenuate SOS, additional time pressure in the reward condition again induced the SOS effect. It has been suggested that screener performance is not linked to salary (Filipczak, 1996; Guzzo, Jette, & Katzell, 1985), but regardless of the accuracy of this claim, any benefits derived from increased motivation in the form of performance-linked compensation may be offset by situations of increased time pressure (e.g., a holiday rush of passengers). The potentially significant factor of time pressure should therefore be an important consideration in the training of airport security screeners and in constructing the environment in which searches are conducted (for instance, obscuring screener awareness of passenger line length).

It is also noteworthy that when we utilized photographic stimuli to partly replicate the perceptual appearance of airport X-ray images, we no longer observed SOS.

We propose that the relatively easy parsing and recognition of our photographic targets may have enabled participants to adequately search each display even after finding a very obvious high-salience target. By reducing decision-making components of visual search (i.e., how much participants dwell on a possible target), extra attentional resources may be available to conduct a better and exhaustive search of the rest of the image. These data suggest a potentially optimistic consequence for airport security, wherein targets are likely to have a high spatial frequency and are more easily parsed than, say, a subtle contrast difference indicative of a pulmonary nodule in a radiological exam. However, caution must be taken in interpreting these lab-based results of untrained participants, considering that our stimulus set was comprised of a finite number of recognizable photographs instead of the infinitely heterogeneous target possibilities which exist in real airport screening.

4.7.5 Conclusion

Our primary goal in this paper has been to interrogate the phenomenon of SOS in a series of controlled laboratory studies to simultaneously inform cognitive psychology, radiology, and real-world searches akin to airport baggage screening. One of the most important implications from this work is that we reveal the first evidence for SOS in a non-medical search. By revealing that SOS can arise for non-experts and in standard cognitive psychology search paradigms, we illustrate some of the factors critical to studying and understanding errors in multi-target search. SOS is a complex,

generalized effect and the only way to reduce its impact is to carefully delineate its variety of underlying causes.

4.8 Appendix: Response time data

One mechanistic explanation for SOS is a 'truncated search,' where participants prematurely end their search after finding a target (Samuel et al., 1995). In the current experiments, participants click on each target they find and then click a button labeled "DONE" once they have decided to terminate their search. Thus, we can ask whether or not participants exhibit a truncated search by comparing the time taken to click "DONE" for high-salience single target trials in which the target was correctly detected and the no-target trials without false alarms. After successfully finding a high-salience target, are participants 'satisfied' and quicker to terminate their search? Although comparing these data is complicated by the fact that one involves an extra mouse-click and that moving a mouse and making a click is less precise of a measure than a keypress, they reveal little evidence of a truncated search; only Experiments 6 and 8 had significantly faster responses for high-salience single target trials than no-target trials (see Table 2 for response time data).

Table 2: Response time data for SOS Experiments 1-13.

Expt	Experimental Parameters			Response Time in Seconds (with SD)			
	Stimuli	Time Limit	Prevalence (High:Low)	No Target	High single	Low single	Dual target
1	Ts & Ls	15s	1:1	9.02 (1.49)	9.43 (1.35)	9.51 (1.38)	8.52 (1.23)
2	Ts & Ls	15s	2:1	9.65 (0.95)	9.42 1.20	9.58 0.86	8.10 0.72
3	Ts & Ls	15s	3:1	9.93 (0.89)	9.53 (0.64)	10.01 (0.70)	8.55 (0.66)
4	Pictures	15s	1:1	8.13 (1.09)	8.73 (1.56)	9.04 (1.43)	7.49 (1.22)
5	Pictures	15s	9:1	8.01 (1.63)	8.12 (1.74)	8.67 (1.90)	6.74 (1.25)
6	Ts & Ls	30s	2:1	10.45 (2.64)	9.39 (2.03)	9.83 (1.96)	8.05 (1.40)
7	Ts & Ls	30s	3:1	10.97 2.87	10.23 2.09	10.72 2.26	8.99 (1.58)
8	Ts & Ls	30s	6:1	11.05 (1.65)	10.38 (1.32)	10.88 (1.65)	8.83 (0.78)
9	Ts & Ls	30s	1:1 (Low-single:dual)	12.47 3.36	-	12.19 (3.13)	10.05 (1.73)
10	Ts & Ls	30s	4:1 (Low-single:dual)	11.99 (2.06)	-	12.02 (1.35)	10.61 (0.83)
11	Ts & Ls	30s (Instructions)	3:1	9.99 (2.53)	9.66 (1.93)	9.57 (1.68)	8.29 (1.48)
12	Ts & Ls	Luggage Line (variable rate)	3:1	10.15 (3.57)	9.46 (2.63)	10.04 (3.23)	7.02 (1.45)
13	Ts & Ls	Luggage Line (constant rate)	3:1	11.77 (4.22)	10.95 (3.08)	11.12 (2.94)	8.65 (1.40)

Chapter 5. Conclusion

Visual search tasks are a perpetual aspect of daily life in visually healthy humans and other species. Search has been studied for decades as a window into the cognition of visual attention because it taps all critical stages between raw sensory perception and final executive processing and motor response behavior. Accordingly, the research reported herein examined visual search both for its pervasiveness in regular life as well as its ability to inform basic cognitive psychology. The aspects of search studied here— expectation, experience, and environment—were chosen specifically because of the potential to directly inform critical real-world searches with the express goal of improving safety and efficiency.

The role of expectation in search was explored in Chapter 1 by utilizing a rare target search paradigm. By manipulating the frequency with which targets appear, searchers may gradually shift expectation about the contents of any given display. Early work suggested that this generated a drastic perceptual failure (Wolfe et al., 2005) in the form a vast increase in miss errors. We further explored the problem by replicating the original design, and although we found a similar increase in miss rates, it was revealed that many participants in fact knew they had made “some errors.” Importantly, a drastic increase in error is only generalizable to the real world if these errors are “true errors” (i.e., not knowingly made), which led to the manipulation presented in the present work: the option to correct mistakes. This simple change eliminated a vast proportion of

“low prevalence errors,” and additional analyses indicated that errors arose when participants were speeding their responses (Fleck & Mitroff, 2007).

It is critical to note that these data do not suggest that there is no such thing as a “low prevalence effect.” Instead, the results emphasize what we believe to be the main consequence of target rarity: responses speed up, often at the expense of accuracy. Faster responding led participants to adopt a prepotent motor response of “target absent,” and in fact many of the errors stemmed directly from an inadvertent press of this response. While the option to correct eliminated many miss errors in this study, this is by no means a generalizable “fix” for all tasks of low prevalence. In this experiment, the simplicity of the stimuli or over-familiarity of participants with the targets and distractors may have enabled very rapid recognition processes that facilitated the correction of many errors, but in more complex tasks that may be typical of real-world searches, the low prevalence effect of speeding responses may lead to situations in which a target is never fixated at all. In such a case, a participant would have no reason to “correct” the prepotent motor response.

The speed-accuracy tradeoff represents a rather parsimonious and perhaps unsurprising account of rare target search performance, but other research has emphasized the contribution of a criterion shift during low prevalence search to the increase in errors (e.g., Wolfe et al., 2007; Van Wert et al., 2009). This model suggests that as low prevalence search proceeds, observers begin to require more and more evidence

to recognize a target rather than reject it as a distractor. The model also suggests that observers may attempt to match the raw number of misses and false alarms across the course of the task (Wolfe et al., 2007). Our data do not disconfirm this model, and explicit modeling of decision parameters (e.g., Madden et al., in revision) will likely be necessary to make stronger conclusions. It remains an important theoretical task to deconvolve the differential contributions of these two related accounts.

Regardless of how low prevalence errors arise, it is helpful to consider individual differences that modulate performance of rare target search. Even if the option to correct eliminates some proportion of prevalence-related errors, establishing which individual experiences offset rare target errors may inform real-world training and further illuminate the low prevalence effect.

In Chapter 2, we explored the effect of video game experience on rare target search. We approached the issue with two key pieces of information: 1) video game players (VGPs) outperform non-video game players (NVGPs) on a wide range of tasks of attention and perception, and 2) VGPs typically respond much faster than NVGPs. With respect to rare target search, these points are important because, as was shown in Chapter 1, faster responding in this particular task is directly linked to lower performance. We wished to know whether VGPs would continue to outperform NVGPs at rare target search, even at faster speeds, which would support a low-level advantage over NVGPs as has been previously proposed (e.g., Green & Bavelier, 2006c).

Even though VGPs indeed performed significantly more accurately than NVGPs on a rare target search, we were surprised by the source of this advantage: slower speeds. In a rare example of VGPs performing more slowly than NVGPs, game players took advantage of the self-paced design (i.e., participants could spend as little or as much time as they wanted on any given trial) to spend extra time on the more difficult low prevalence trials than on the easier high prevalence trials, consequently detecting rare targets at near-ceiling accuracy. These data again emphasize the speed-accuracy tradeoff described in Chapter 2, but more importantly they indicate that experience can indeed modulate rare target search. Specifically, the drastic slowdown suggests that the experience with video games may result in a strategic or motivational difference in approach to the task, rather than a strictly low-level resource advantage that previous research implicated (see Green & Bavlier, 2006c). Additionally, the greater variation in response times for VGPs than for NVGPs may indicate more of a metacognitive ability to *detect* the situations in which slower speeds are necessary for optimal performance.

Together, the results of Chapter 2 add to previous work (Rosser et al., 2007) suggesting that video game experience might positively modulate real-world performance. Although we used a simple search task, we modeled the stimuli and response parameters after the X-ray screening procedure found in airport security. The data suggest that it may be worth considering extensive video game play as helpful experience when recruiting or training for critical real-world tasks such as airport

security or radiological examinations. Although we did not establish the precise mechanisms responsible for the improved performance, the slower response times in VGP's suggest that a higher-level effect, such as strategy or motivation, generates the slower behavior in response to the more difficult demands of rare target search.

The final set of experiments presented herein explored the role of extra targets in the search environment. The detection of one target in radiology has been shown to have a detrimental effect on the detection of additional targets in the same case or image, a phenomenon known as "satisfaction of search" (SOS; e.g., Tuddenham, 1962). However, outside of the radiological domain, this effect has largely been underexplored, yet it is easy to witness many instances of multiple target search in real-world tasks, including airport security, cytological examinations, and a variety of other professional inspections. Here we examined the generality of SOS: Does it arise outside of radiology? Additionally, we continued to explore the role of expectation by manipulating the frequency of particular targets, in an attempt to understand any potential interactions between expectancy of a certain target type and the presence or absence of another target type.

Across a series of thirteen experiments we manipulated time pressure, discriminability, salience, frequency, instructions, and reward. We found that SOS indeed generalizes outside of radiology but is not a consequence of just any multiple

target search. Rather, the SOS effect is sensitive, to varying degrees, to particular pressures and the interplay between them.

Specifically, we found that salience differences in a multiple target search (target A is easier to see than target B) are not sufficient to induce SOS, but that when the easier target is also more common than the more difficult target, the rarer target suffers from an SOS effect. This finding carries potentially important implications for airport security, in which recent protocol changes now demand screening for liquids, which are far more frequent and salient than some other target categories, such as hidden weapons. Interestingly, the effect was minimized when participants were alerted in advance about relative frequencies of target types, suggesting a metacognitive ability to ward off expectation effects.

We also established that this interaction between salience and frequency no longer generates SOS when target items are more easily recognized, suggesting that a per-item decision making process may be a necessary precursor to the SOS effect. Consequently, the higher spatial frequency of targets in airport security (as opposed to the subtle contrast differences indicative of targets in chest radiography, for example), may lead to less SOS than is seen in radiological contexts.

Next, we revealed that time pressure can exacerbate the SOS effect. When participants were rushed to complete the task, either by imposing a 15 second time limit or by rewarding faster responses, an SOS effect was stronger than when participants had

twice as much time or less reward pressure. These data emphasize that across all domains—radiology, cytology, airport security, and so on—although most tasks are self-paced, it is nevertheless important to minimize the time pressures that are associated with such searches.

A remaining theoretical question is how video game expertise modulates multiple target search. As Chapter 3 indicated, video game expertise is typified by a metacognitive assessment of the difficulty of a task and a corresponding behavioral adjustment to maximize performance, suggesting that gamers might effortfully offset the costs of SOS by exhaustively searching each display after finding a target. Yet, the relatively robust SOS effect speaks to a generalized search heuristic, and if the findings in Chapter 3 reflect more of a strategic benefit rather than a motivational difference, gamers might more quickly adopt the search heuristic which indeed is on some level a very adaptive technique to perform adequately and efficiently. Further research is necessary to gauge the effect of video game experience on multiple target search, and the work may help distinguish between the “higher level” components of motivation and strategy as proposed in Chapter 3.

The influences on visual search explored in the research presented herein, including expectation, experience, and environmental factors, all have been shown to significantly impact the execution of visual search tasks. These higher order factors directly influence the response times and accurate detection of targets. By moving

beyond display-level parameters such as features, set size, and search efficiency, we offer here data that can directly inform the design and implementation of critical visual search tasks. Ideally, these results can help shape the training, recruitment, and protocols of real-world search tasks such as those found in radiology, cytology, and airport security screening.

References

- Ashman, C. J., Yu, J. S., & Wolfman, D. Satisfaction of search in osteoradiology. (2000). *American Journal of Roentgenology*, 177, 252-253.
- Berbaum, K. S., El-Khoury, G. Y., Franken Jr., E. A., Kuehn, D. M., Meis, D. M., Dorfman, D. D., et al. (1994). Missed fractures resulting from satisfaction of search effect. *Emergency Radiology*, 1, 242-249.
- Berbaum, K. S., El-Khoury, G. Y., Ohashi, K., Schartz, K. M., Caldwell, R. T., Madsen, M. T., et al. (2007). Satisfaction of search in multi-trauma patients: severity of detected fractures. *Academic Radiology*, 14, 711-722.
- Berbaum, K. S., Franken Jr., E. A., Caldwell, R. T., & Schartz, K. M. (2006). Can a checklist reduce SOS errors in chest radiography? *Academic Radiology*, 13, 296-304.
- Berbaum, K. S., Franken Jr., E. A., Dorfman, D. D., Caldwell, R. T., & Krupinski, E. A. (2000). Role of faulty decision making in the satisfaction of search effect in chest radiography. *Academic Radiology*, 7, 1098-1106.
- Berbaum, K. S., Franken Jr., E. A., Dorfman, D. D., Miller, E. M., Krupinski, E. A., Kreinbring, K., et al. (1996). The cause of satisfaction of search effects in contrast studies of the abdomen. *Academic Radiology*, 3, 815-826.
- Berbaum, K. S., Franken Jr., E. A., Dorfman, D. D., Miller, E. M., Caldwell, R. T., Kuehn, D. M., et al. (1998). Role of faulty visual search in the satisfaction of search effect in chest radiography. *Academic Radiology*, 5, 9-19.
- Berbaum, K. S., Franken Jr., E. A., Dorfman, D. D., Rooholamini, S. A., Kathol, M. H., Barloon, T. J., et al. (1990). Satisfaction of search in diagnostic radiology. *Investigative Radiology*, 25, 133-140.
- Berlin, L. (1994). Reporting the "missed" radiologic diagnosis: medicolegal and ethical considerations. *Radiology*, 192, 183-187.
- Boot, W. R., Kramer, A. F., Fabiani, M., Gratton, G., Simons, D. J., Wan X. I., et al. (2006). The effects of video game playing on perceptual and cognitive abilities. *Journal of Vision*, 6, 942.
- Bowditch, R. (1996). Patterns found in false negative cervical cytology. *Cytoletter*, 3, 22-25.

- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 443-446.
- Broadbent, D. E., & Gregory, M. (1963). Vigilance considered as a statistical decision. *British Journal of Psychology, 44*, 309-323.
- Broadbent, D. E., & Gregory, M. (1965). Effects of noise and of signal rate upon vigilance analysed by means of decision theory. *Human Factors, 7*, 155-162.
- Buck, L. (1966). Reaction time as a measure of perceptual vigilance. *Psychological Bulletin, 65*, 291-308.
- Bundesen, C., & Pedersen, L. F. (1983). Color segregation and visual search. *Perception and Psychophysics, 33*, 487-493.
- Carter, R. C. (1982). Visual search with color. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 127-136.
- Castel, A. D., Pratt, J., & Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta Psychologica, 119*, 217-230.
- Chapanis, A. (1999). *The Chapanis Chronicles: 50 Years of Human Factors Research, Education, and Design*. Santa Barbara, CA: Aegean.
- Chun, M. M., & Jiang, Y. (1998). Contextual cuing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*, 28-71.
- Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology, 30*, 39-78.
- Davies, D. R., Shackleton, V. J., & Parasuraman, R. (1983). Monotony and boredom. In G. R. J. Hockey (Ed.), *Stress and Fatigue in Human Performance*. (pp. 1-32). New York: Wiley.
- De Lisi, R., & Cammarano, D. M. (1996). Computer experience and gender differences in undergraduate mental rotation performance. *Computers in Human Behavior, 12*, 351-361.
- DeMay, R. M. (1997). Common problems in Papanicolaou smear interpretation. *Archives of Pathology and Laboratory Medicine, 121*, 229-238.

- Di Lollo, V., Kawahara, J., Zuvic, S. M., & Visser, T. A. W. (2001). The preattentive emperor has no clothes: A dynamic redressing. *Journal of Experimental Psychology: General*, 130, 479-492.
- Dick, M., Ullman, S., & Sagi, D. (1987). Parallel and serial processes in motion detection. *Science* 237, 400-402.
- Dorval, M., & Pepin, M. (1986). Effect of playing a video game on a measure of spatial visualization. *Perceptual Motor Skills*, 62, 159-162.
- Drew, D., & Waters, J. (1986). Video games: Utilization of a novel strategy to improve perceptual motor skills and cognitive functioning in the non-institutionalized elderly. *Cognitive Rehabilitation* 4, 26-31.
- Drury, C. G., & Addison, J. L. (1973). An industrial study of the effects of feedback and fault density on inspection performance. *Ergonomics*, 16, 159-169.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501-517.
- Duncan, J. (1989). Boundary conditions on parallel processing in human vision. *Perception*, 18, 457-469.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433-458.
- Egeth, H. & Smith, E. E. (1967). On the nature of errors in a choice reaction task. *Psychonomic Science*, 8, 345-346.
- Eggin, T. K., & Feinstein, A. R. (1996). Context bias: a problem in diagnostic radiology. *The Journal of the American Medical Association*, 276, 1752-1755.
- Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, 123, 161-177.
- Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Current Biology*, 6, 292-297.
- Filipczak, B. (1996). Can't buy me love. *Training*, 33, 29-35.

- Fiore, S. M., Scielzo, S., & Jentsch, F. (2004). Stimulus competition during perceptual learning: Training and aptitude considerations in the X-ray security screening process. *International Journal of Cognitive Technology, 9*, 34-39.
- Fiorentini, A., & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature, 287*, 43-44.
- Fitts, P.M. (1947). Psychological research on equipment design (Research Report 19). Washington, DC: U.S. Army Air Forces Aviation Psychology Program.
- Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science, 18*, 943-947.
- Folk, C., 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*, 1030-1044.
- Foster, D. H., & Ward, P. H. (1991). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proceedings of the Royal Society of London, Series B. Biological Sciences, 243*, 75-81.
- Franken, E. A., Berbaum, K. S., Lu, C. H., Kannam, S., Dorfman, D. D., Warnock, N. G., et al. (1994). Satisfaction of search in the detection of plain-film abnormalities in abdominal contrast studies. *Investigative Radiology, 29*, 403-409.
- Gagnon, D. (1985). Videogame and spatial skills: an explanatory study. *Educational Communication and Technology Journal, 33*, 263-275.
- Gale, A. G., Mugglestone, M. D., Purdy, K. J., & McClumpha, A. (2000). Is airport baggage inspection just another medical image? *Proceedings of the SPIE, 3981*, 184-192.
- Gerritsen, C., Frischen, A., Blake, A., Smilek, D., & Eastwood, J. D. (2008). Visual search is not blind to emotion. *Perception & Psychophysics, 70*, 1047-1059.
- Gibson, B. S., Li, L., Skow, E., Brown, K., & Cooke, L. (2000). Searching for one versus two identical targets: When visual search has a memory. *Psychological Science, 11*, 324-327.
- Goldstone, R. L. (1995). Effects of categorization on color perception. *Psychological Science, 6*, 298-304.

- Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585-612.
- Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors, 36*, 387-405.
- Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*, 534-537.
- Green, C. S., & Bavelier, D. (2006a). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1465-1478.
- Green, C. S., & Bavelier, D. (2006b). Enumeration versus multiple object tracking: the case of action video game players. *Cognition, 101*, 217-245.
- Green, C. S., & Bavelier, D. (2006c). The cognitive neuroscience of video games. In P. Messaris & L. Humphreys (Eds.), *Digital Media: Transformations in Human Communication*. (pp. 211-224). New York: Peter Lang.
- Green, C. S., & Bavelier, D. (2007). Action video game experience alters the spatial resolution of vision. *Psychological Science, 18*, 88-94.
- Green, D. M., & Swets J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.
- Greenfield, P. M., DeWinstanley, P., Kilpatrick, H., & Kaye, D. (1994). Action video games and informal education: Effects on strategies for dividing visual attention. *Journal of Applied Developmental Psychology, 15*, 105-123.
- Griffith, J. L., Voloschin, P., Gibb, G. D., & Bailey, J. R. (1983). Differences in eye-hand motor coordination of video-game users and non-users. *Perceptual and Motor Skills, 57*, 155-158.
- Gur, D., Rockette, H. E., Armfield, D. R., Blachar, A., Bogan, J. K., Brancatelli, G., et al. (2003). The prevalence effect in a laboratory environment. *Radiology, 228*, 10-14.
- Gur, D., Rockette, H. E., Warfel, T., Lacomis, J. M., & Fuhrman, C. R. (2003). From the laboratory to the clinic: The "Prevalence Effect." *Academic Radiology, 10*, 1324-1326.

- Guzzo, R. A., Jette, R. D., & Katzell, R. A. (1985). The effects of psychologically based intervention programs on worker productivity: A meta-analysis. *Personnel Psychology, 38*, 275-291.
- Hanson, C. H., & Hanson, R. D. (1988). Finding the face in the crowd: The anger superiority effect. *Journal of Personality and Social Psychology, 54*, 917-924.
- Hillstrom, A. P. (2000). Repetition effects in visual search. *Perception and Psychophysics, 62*, 800-817.
- Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394*, 575-577.
- James, W. *The Principles of Psychology, v.1*. (1890). (pp. 402-403). New York: Henry Holt.
- Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 179-219.
- Karni A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature, 365*, 250-252.
- Kinard, R. E., Orrison, W. W., & Brogdon, B. G. (1986). The value of a worksheet in reporting body-CT examinations. *American Journal of Roentgenology, 147*, 848-849.
- Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science, 10*, 346-352.
- Körner, C., & Gilchrist, I. D. (2008). Memory processes in multiple-target visual search. *Psychological Research, 72*, 99-105.
- Kristjansson, A. (2000). In search of remembrance: Evidence for memory in visual search. *Psychological Science, 11*, 328-332.
- Kundel, H. L. (1989). Perception errors in chest radiography. *Seminars in Respiratory and Critical Care Medicine, 10*, 203-210.
- Kundel, H. L., Nodine, C. F., & Carmody, D. (1978). Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Investigative Radiology, 13*, 175-181.
- Lanagan-Leitzel, L. K., & Moore, C. M. (2008). Novice and expert performance on a computerized lifeguarding task. *Journal of Vision, 8*, 318.

- Leber, A. B., & Egeth, H. E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, *13*, 132-138.
- Li, F., Li, H., Yan, J. H., Gao, H. H., Chen, A., & Lin, C. (unpublished). Appropriate responding reduces missing errors in visual search.
- Lintern, G., & Kennedy, R. S. (1984). Video game as a covariate for carrier landing research. *Perceptual and Motor Skills*, *58*, 167-172.
- MacDonald, C. J., & Meck, W. H. (2004). Systems-level integration of interval timing and reaction time. *Neuroscience and Biobehavioral Reviews*, *28*, 747-769.
- Mackworth, N. H. (1950). Researches on the measurement of human performance. *Medical Research Council Special Report, London, Series 268*.
- Madden, D. J., & Langley, L. K. (2003). Age-related changes in selective attention and perceptual load during visual search. *Psychology and Aging*, *18*, 54-67.
- Madden, D. J., Mitroff, S. R., Shepler, A. M., Fleck, M. S., Costello, M. C., & Voss, A. (in revision). Rare target search: Diffusion model analysis and effects of adult age.
- Maljkovic, V., & Nakayama, K. (1994). Priming of popout: I. Role of features. *Memory & Cognition*, *22*, 657-672.
- Manning, D. J., Ethell, S. C., & Donovan, T. (2004). Detection or decision errors? Missed lung cancer from posteroanterior chest radiograph. *British Journal of Radiology*, *77*, 231-235.
- McCarley, J. S., & Carruth, D. W. (2004). Oculomotor scanning and target recognition in luggage X-ray screening. *Scanning and Recognition*, *9*, 26-29.
- McCarley, J. S., Kramer, A. F., Wickens, C. D., Vidoni, E. D. & Boot, W. R. (2004). Visual skills in airport-security screening. *Psychological Science*, *15*, 302-306.
- Melcher, D., Papathomas, T. V., & Vidnyánszky, Z. (2005). Implicit attentional selection of bound visual features. *Neuron*, *46*, 723-729.
- Menneer, T., Barrett, D. J. K., Phillips, L., Donnelly, N., & Cave, K. R. (2006). Costs in searching for two targets: dividing search across target types could improve airport security screening. *Applied Cognitive Psychology*, *21*, 915-932.
- Metlay, W., Sokoloff, M., & Kaplan, I. T. (1970). Visual search for multiple targets. *Journal of Experimental Psychology*, *85*, 148-150.

- Most, S. B., Scholl, B. J., Clifford, E., & Simons, D. J. (2005). What you see is what you set: Sustained inattentive blindness and the capture of awareness. *Psychological Review*, *112*, 217-242.
- Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be seen: the contribution of similarity and selective ignoring to sustained inattentive blindness. *Psychological Science*, *12*, 9-17.
- Münsterberg, H. (1913). *Psychology and Industrial Efficiency*. New York: Houghton.
- Nagy, A. L., & Thomas, G. (2003). Distractor heterogeneity, attention, and color in visual search tasks. *Vision Research*, *43*, 1541-1552.
- Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.
- Obuchowski, N. A. (2005). One less bias to worry about. *Radiology*, *232*, 302.
- O'Riordan, M. A. F., Plaisted, K. C., Driver, J., & Baron-Cohen, S. (2001). Superior visual search in autism. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 719-730.
- Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*, 1227-1268.
- Pastore, R. E., Crawley, E. J., Berens, M. S., & Skelly, M. A. (2003). "Nonparametric" A' and other modern misconceptions about signal detection theory. *Psychonomic Bulletin & Review*, *10*, 556-569.
- Plaisted, K. C., O'Riordan, M. A. F., & Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: a research note. *Journal of Child Psychology and Psychiatry*, *39*, 777-783.
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds.), *Attention and Performance X*. (pp. 531-556). Hillsdale, NJ: Erlbaum.
- Quinlan, P. T., & Humphreys, G. W. (1987). Visual search for targets defined by combinations of color, shape, and size: An examination of the task constraints on feature and conjunction searches. *Perception and Psychophysics*, *41*, 455-472.
- Rabbitt, P. M. A. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, *71*, 264-272.

- Rauschenberger, R., & Yantis, S. (2001). Masking unveils pre-amodal completion representation in visual search. *Nature*, *410*, 369-372.
- Remington, R. W., Johnston, J. C., & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception and Psychophysics*, *51*, 279-290.
- Renfrew, D. L., Franken, E. A., Berbaum, K. S., Weigelt, F. H., & Abu-Yousef, M. M. (1992). Error in radiology: classification and lessons in 182 cases presented at a problem case conference. *Radiology*, *183*, 145-150.
- Rensink, R. A. (2000). Visual search for change: A probe into the nature of attentional processing. *Visual Cognition*, *7*, 345-376.
- Rensink, R. A., & Enns, J. T. (1998). Early completion of occluded objects. *Vision Research*, *28*, 169-184.
- Rich, A. N., Kunar, M. A., Van Wert, M. J., Hidalgo-Sotelo, B., Horowitz, T. S., & Wolfe, J. M. (2008). Why do we miss rare targets? Exploring the boundaries of the low prevalence effect. *Journal of Vision*, *8*, 1-17.
- Rose, M. (1975). *Industrial Behavior*. London: Allen Lane.
- Rosenberg, B. H., Landsittel, D. S., & Averch, T. D. (2005). Can video games be used to predict or improve laparoscopic skills. *Journal of Endourology*, *19*, 372-376.
- Rosser, Jr, J. C., Lynch, P. J., Cuddihy, L., Gentile, D. A., Klonsky, J., & Merrell, R. (2007). The impact of video games on training surgeons in the 21st century. *Archives of Surgery*, *142*, 181-186.
- Rubenstein, J. (2001). (Ed.) *Test and evaluation plan: x-ray image screener selection test*. Washington, DC: Office of Aviation Research.
- Samuel, S., Kundel, H. L., Nodine, C. F., & Toto, L. C. (1995). Mechanisms of satisfaction of search: Eye position recordings in the reading of chest radiographs. *Radiology*, *194*, 895-902.
- Schwaninger, A. (2005). Increasing efficiency in airport security screening. *WIT Transactions on the Built Environment*, *82*, 405-416.
- Schwaninger, A., Hardmeier, D., & Hofer, F. (2005). Aviation security screeners visual abilities & visual knowledge measurement. *IEEE Aerospace and Electronic Systems*, *20*, 29-35.

- Sekuler, A. B., Palmer, S. E., & Flynn, C. (1994). Local and global processes in visual completion. *Psychological Science, 5*, 260-267.
- Simon, H. (1976). *Administrative Behavior (3rd ed.)*. New York: The Free Press
- Sims, V. K., & Mayer, R. E. (2002). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology, 16*, 97-115.
- Smith, J. D., Redford, J. S., Washburn, D. A., & Tagliatela, L. A. (2005). Specific-token effects in screening tasks: Possible implications for aviation security. *Journal of Experimental Psychology: Learning, Memory & Cognition, 31*, 1171-1185.
- Subrahmanyam, K., & Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology, 15*, 13-32.
- Takeda, Y. (2004). Search for multiple targets: Evidence for memory-based control of attention. *Psychonomic Bulletin & Review, 11*, 71-76.
- Townsend, J. T. (1990). Serial and parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science, 1*, 46-54.
- Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 194-214.
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology, 40*, 201-237.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97-136.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological Review, 95*, 15-48.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14*, 107-141.
- Trick, L. M., Jaspers-Fayer, F., & Sethi, N. (2005). Multiple-object tracking in children: The "Catch the Spies" task. *Cognition, 20*, 373-387.

- Tuddenham, W. J. (1962). Visual search, image organization, and reader error in roentgen diagnosis: Studies of the psycho-physiology of roentgen image perception. *Radiology*, 78, 694-704.
- Van Wert, M. J., Horowitz, T. S., & Wolfe, J. M. (2009). Even in correctable search, some types of rare targets are frequently missed. *Attention, Perception, & Psychophysics*, 71, 541-553.
- Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception and Psychophysics*, 56, 495-500.
- Wiegmann, D., McCarley, J. S., Kramer, A. F., & Wickens, C. D. (2006). Age and automation interact to influence performance of a simulated luggage screening task. *Aviation, Space, and Environmental Medicine*, 77, 825-831.
- Wilbur, D. C. (1997). False negatives in focused rescreening of Papanicolaou smears: How frequently are 'abnormal' cells detected in retrospective review of smears preceding cancer or high-grade intraepithelial neoplasia? *Archives of Pathology & Laboratory Medicine*, 121, 273-276.
- Wixted, J. T., & Stretch, V. (2000). The case against a criterion-shift account of false memory. *Psychological Review*, 107, 368-376.
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention*. (pp. 13-73). East Sussex, UK: Psychology Press.
- Wolfe, J. M. (2001). Asymmetries in visual search: an introduction. *Perception & Psychophysics*, 63, 381-389.
- Wolfe, J. M. (2007). Guided Search 4.0: Current Progress with a model of visual search. In W. Gray (Ed.), *Integrated Models of Cognitive Systems*. (pp. 99-119). New York: Oxford.
- Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature*, 435, 439-440.
- Wolfe, J. M., Horowitz, T. S., Van Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology: General*, 136, 623-638.
- Wyatt, S., & Langdon, J. N. (1932). Inspection processes in industry. *IFRB Report*, No. 63. London: HMSO.

Wyatt, S., Langdon, J. N., & Stock, F. G. L. (1937). Fatigue and boredom in repetitive work. *IFRB Report, No. 77*. London: HMSO.

Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention*. (pp. 223-256). East Sussex, UK: Psychology Press.

Biography

Mathias (Mat) Samuel Fleck was born March 26, 1978 in Chicago, Illinois and grew up in nearby Beecher, Illinois for 10 years before moving to Chesterton, Indiana. He attended Chesterton High School and was a state and national finalist in the school's national-ranked Speech & Debate program. Mat matriculated at the University of Chicago in the Early Entrance program after three years of high school, receiving a Bachelor of Arts degree in Psychology in June, 1999. After college, he co-founded FlexGames, Inc. in Long Beach, CA to design and develop computer and web games. In 2003, Mat began the Ph.D program in the Department of Psychology & Neuroscience at Duke University in Durham, NC in the lab of Dr. Roberto Cabeza, using functional magnetic resonance imaging (fMRI) to study episodic memory. In 2006, Mat joined the visual cognition lab of Dr. Stephen Mitroff and in 2008 received a Ruth L. Kirschstein National Research Service Award (NRSA) to study influences on visual search performance.

Education

Duke University, Durham, NC

Ph.D., Psychology & Neuroscience, May 2009

Dissertation: Effects of Expectation, Experience, and Environment on Visual Search

Committee: Stephen Mitroff (chair), Amy Needham, David Madden, Elizabeth Marsh

University of Chicago, Chicago, IL

B.A. Psychology, June 1999

Publications (in peer reviewed journals)

- Mitroff, S. R., Arita, J. T., & **Fleck, M. S.** (2009). Staying in bounds: Contextual constraints on object file coherence. *Visual Cognition, 17*, 195-211.
- Davis, S., Dennis, N., Daselaar, S., **Fleck, M. S.**, & Cabeza, R. (2008). Qué PASA? The posterior-anterior shift in aging. *Cerebral Cortex, 18*, 1201-1209.
- Fleck, M. S.**, & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science, 18*, 943-947.
- Fleck, M. S.**, Daselaar, S. M., Dobbins, I. G., & Cabeza, R. (2006). Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and non-memory tasks. *Cerebral Cortex, 16*, 1623-1630.
- Daselaar, S. M., **Fleck, M. S.**, Prince, S., & Cabeza, R. (2006). The medial temporal lobe distinguishes old from new independently of consciousness. *The Journal of Neuroscience, 26*, 5835-5839.
- Daselaar, S. M., **Fleck, M. S.**, & Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: recollection, familiarity, and novelty. *Journal of Neurophysiology, 96*, 1902-1911.
- Daselaar, S. M., **Fleck, M. S.**, Dobbins, I. G., Madden, D. J., & Cabeza, R. (2006). Effects of healthy aging on hippocampal and rhinal memory functions: An event-related fMRI study. *Cerebral Cortex, 16*, 1771-1782.

Publications (submitted or in preparation)

- Fleck, M. S.**, Samei, E., & Mitroff, S. R. (submitted). Generalized 'satisfaction of search': Adverse influences on dual-target search.
- Fleck, M. S.**, & Mitroff, S. R. (in preparation). Video game players excel at rare target visual search.
- Madden, D. J., Mitroff, S. R., Shepler, A. M., **Fleck, M. S.**, Costello, M. C., & Voss, A. (in revision). Rare target search: Diffusion model analysis and effects of adult age.
- Clark, K., **Fleck, M. S.**, & Mitroff, S. R. (in preparation). Effects of videogame expertise on change detection abilities.

Awards

- | | |
|-----------|---|
| 2008-2009 | NIH Ruth L. Kirschstein NRSA Predoctoral Fellowship |
| 2007-2008 | Duke Vertical Integration Mentorship Program |
| 2007 | Conference Travel Fellowship, Duke University |
| 2005 | National Science Foundation Predoctoral Fellowship, Honorable Mention |