Overcoming Hurdles in Translating Visual Search Research Between the Lab and the Field

Kait Clark, Matthew S. Cain, Stephen H. Adamo, & Stephen R. Mitroff

Center for Cognitive Neuroscience, Department of Psychology and Neuroscience

Duke University

Chapter for the 59th Nebraska Symposium on Motivation

Address for correspondence	:	Kait Clark Center for Cognitive Neuroscience Duke University LSRC Building, Box 90999 450 Research Drive Durham, NC 27708
Email	:	kait.clark@duke.edu
Phone/Fax	:	(919) 668-6144

1. General introduction

Imagine two hypothetical individuals, Tyler and Olivia. Tyler is an undergraduate at Duke University. He is taking an Introductory Psychology course that requires him to participate in three experimental psychology studies. For one of the required experiments, Tyler selects a "visual search" study in a cognitive psychology laboratory and signs up at a time immediately after his morning Linear Algebra class. Tyler arrives at the lab at 11am and fills out the necessary consent forms before being led into a dimly lit testing room where he is instructed to search for T-shaped targets on a computer screen among L-shaped distractor items and to indicate whether a T is present or absent on each trial. Tyler is bored by the time the practice segment is complete and begins to muddle through the task, exerting the minimal effort required. While he is sure some graduate student really cares about how he performs, he is not too concerned. After checking his email on his phone for the 25th time, he finally nears the end of this exercise. He speeds up as the end of the hour approaches; his stomach is beginning to rumble, and he's getting hungry. He knows that the faster he can get through the remainder of the trials, the sooner he'll be able to eat lunch. Tyler completes the task and heads to a university café while the experimenters in the lab examine his data, pooled with data from other participants just like Tyler, with the intentions of drawing conclusions about the nature of human visual search processes.

Olivia is an X-ray operator at the Raleigh-Durham Airport in North Carolina. She has worked with the Transportation Security Administration for six years and is currently a full-time employee who works five days a week. She starts her normal shift at 5am, and

during each shift she works several 30-minute stints at the X-ray machine, searching for contraband that may be hidden in passengers' luggage. Olivia has completed an initial training on optimal search strategies and numerous refresher courses that are designed to make her a better searcher. As her supervisors monitor her performance, she is cognizant of maintaining a certain level of accuracy in order to keep her job. Additionally, Olivia is keenly aware of the consequences of letting a bag with a bomb slip through the cracks. Very few, if any, of the bags Olivia searches contain any actual lethal items, but her job is to remain vigilant and conduct thorough searches on every bag, regardless of the improbability of finding a bomb, a gun, or a knife.

Clearly, our hypothetical individuals, Tyler and Olivia, are conducting very different visual searches, with very different motives, in very different environments. Moreover, the differences highlighted above only scratch the surface of the variability between the nature of inexperienced and expert searchers. Realistically, how can search performance between undergraduates and career searchers be compared, given the drastic differences between these scenarios? Can we conclude anything from Tyler's performance about how Olivia should be performing her job? Does understanding Olivia's performance inform cognitive theories of visual search?

An eventual goal of basic research is translating findings from the **lab** to the **field**[1]; however, researchers often struggle to overcome the inherent differences between the sterile, controlled environment of a research lab and the complex, messy environment of

[1] For the purposes of this chapter, we will use "lab" to refer to visual search experiments conducted by cognitive psychologists with inexperienced searchers in a laboratory setting, and we will use "field" to refer to visual searches conducted as part of normal activities in naturalistic settings that are often done by highly trained "expert" searchers.

the real world. The study of visual search — finding a target amongst distractors — is an excellent paradigm to illustrate this relationship. On the one hand, visual search is a powerful research method for psychologists, as it encompasses several aspects of cognition (e.g., memory, perception, attention). As such, researchers have extensively studied and theorized about the nature of visual search (see Nakayama & Martini, 2010; Eckstein, 2011, for recent reviews). On the other hand, visual searches are regularly conducted in everyday tasks outside of a laboratory setting.

The goal of this chapter is to discuss several of the hurdles encountered when moving between the lab and the field and how they might be overcome. We will begin with a general overview of visual search followed by a brief review of the research history and theories. We will then introduce some relevant applied visual search findings before detailing four primary hurdles that stand in the way of translating search findings between the lab and the field:

1. **Target prevalence**: Is search performance affected by the relative likelihood of a target being present (e.g., do searchers perform worse if targets are rarely present)?
2. **Number of targets and target categories**: Does search performance decline if a searcher is required to search for more than one target in the same image (e.g., multiple fractures in a medical X-ray) or more than one possible target type (e.g., a gun *or* a bomb in a luggage X-ray)?

3. **Motivation and anxiety**: How does the context within which a search is conducted affect performance? Is search performance helped or hindered by added motivation or anxiety?
4. **Level of experience**: Is search performance altered by expertise? How might years of experience alter visual search strategies or abilities?

Overview of visual search

Visual search is the process of finding specific target items within an environment based on particular visual features or semantic information. In its simplest form, visual search could operate via basic pattern matching; for example, detecting a red vertical line in a field of green horizontal lines would solely require invoking a red and/or vertical pattern template. However, even this easy visual search depends on many attentional and perceptual factors, and more complex searches move well beyond basic pattern matching. More common visual searches, both in the lab (e.g., finding a target 'T' amongst distractor 'L's, finding a particular shape amongst variable distractors) and in the field (e.g., finding a tumor in a radiograph, finding keys in a purse), involve an array of cognitive processes. Search involves perception (i.e., processing and interpreting visual features), attention (i.e., allocating resources to the relevant areas of a visual area), and memory (i.e., storing a representation of the target item or items). Thus, search represents a fruitful and exciting area of research.

In the lab, visual search has been used extensively to learn about cognition. For example, search studies have informed theories of basic perception (e.g., Wolfe, Birnkrant,

Kunar, & Horowitz, 2005), the structure of visual short-term memory (e.g., Alvarez & Cavanagh, 2004), and attentional capture (e.g., Yantis & Jonides, 1996; Franconeri, Hollingworth, & Simons, 2005), to name just a few. Beyond using visual search as a powerful tool for understanding cognitive processing, researchers have also focused on search as an experimental paradigm with the goal of understanding how searches are conducted. Over the past several decades, psychological research has made tremendous headway in understanding the processes responsible for performing visual search tasks and the mechanisms that allow for the successful identification of target items. The findings from visual search research have been extensive, and, in turn, the contributions to the scientific community have been invaluable.

Given the relevance of visual search to real-world environments as well, ideally, what is learned from studying search processes in the lab can be applied to searches in the field. Beyond the vital function of search in navigating our everyday lives, the search performance of radiologists, X-ray operators, and many others can be life-or-death critical. As recent technological advancements have allowed for the improvement of screening techniques, additional key advancements lie in understanding the cognitive processes of the searchers themselves, identifying common search errors, and improving the manner in which searches are conducted.

A brief history of visual search research

Early evidence from non-human visual searchers.

While the current era of visual search research is largely laboratory-based, the first investigations of search were focused on its primary goal in the world — survival. Animals engage in survival activities that require visual search, such as finding food, avoiding predators, detecting a potential mate's signs, and locating appropriate shelter. Search was perhaps first scientifically investigated in 1890 by Edward Poulton, a zoologist who was interested in how animals elude predators. Poulton noted that a single species tends to evolve many different appearances, making it more difficult to be detected by predators, a phenomenon known as *cryptic pattern polymorphism*. For instance, a single species of forest moth appears with many different wing patterns, and Poulton noticed that it is more difficult for a bird to search for a multiple kinds of targets simultaneously than to search for a single type. The added difficulty that accompanies a search for multiple kinds of targets is now a well-documented idea in cognitive psychology (e.g., Menneer, Barrett, Phillips, Donnelly, & Cave, 2007) and is especially relevant to current-day X-ray baggage screening at airports, as security officers must search for a number of potentially hazardous items.

Expanding on Poulton's observations, Tinbergen (1960) found that insectivorous birds maximized their rate of detection by confining their searches to only a few prey types at a given time and by focusing on either the most common prey available or those that had been seen most recently. In effect, this research demonstrated that non-human animals are sensitive to the statistics of their environments and are able to quickly adapt to maximize search efficiency, and contemporary work with human searchers has found similar results (e.g., Cain, Vul, Clark, & Mitroff, in press). Pigeon studies have illustrated

that search is specialized for ecologically relevant tasks, as pigeons demonstrate a fantastic ability to find food (e.g., Bond, 1983) and effectively optimize their rate of food discovery. These early studies of search with non-human species have served to both establish the evolutionary basis of search processes and demonstrate the practical nature of visual search. For the remainder of this chapter, we focus on human visual search research that has built upon, and complements, these and other non-human search findings.

Early evidence from human visual searchers.

Speculations about the nature of human visual search—also from an applied angle—began with Bernard Koopman in the 1950s, when he explored theories of search in the context of radar operators locating enemy ships (Koopman, 1956a; 1956b). Tasked by the U.S. Navy to systematically determine the location of enemy ships and lost personnel, he revealed many basic theoretical properties of visual search, such as the distribution of attention and the criteria for termination (Koopman, 1957), that remain fundamentally important for current theories of search (e.g., Chun & Wolfe, 1996).

Cognitive psychologists entered the visual search research arena in the 1960s and 1970s and have played a primary role ever since. Early work (e.g., Neisser 1963, Schneider & Shriffran, 1977) laid the groundwork for two influential theories (see Palmer, Verghese, & Pavel, 2000, for a review): the *feature integration theory* (Treisman & Gelade, 1980) and the *guided search model* (Wolfe, 1998).

While seeking to isolate the fundamental elements of vision, Treisman developed the feature-integration theory (FIT; Treisman & Gelade, 1980), which served as a driving

force of the surge of research in visual search that was soon to follow. The basic idea behind FIT is rooted in Neisser's (1967) original division of visual processing into two distinct stages, but Treisman expanded on the meaning of these stages dramatically. According to the theory, the basic features of items (color, shape, orientation, etc.) are first processed effortlessly and automatically in the early stages of vision, in separate, spatially organized maps. Next, directed attention is required during the "attention" stage in order to successfully bind the separate features into integrated object percepts (Treisman, 1998). Finally, a subset of these items is selected for further processing.

FIT allows for the dissociation of two types of searches, often categorized as "parallel" and "serial." *Parallel* search occurs when all items in a search array are assessed simultaneously, as the target item is different from all distractor items on at least one dimension, so it simply "pops out" at the observer (e.g., a red vertical line amongst green horizontal lines). In parallel searches, increasing the set size (number of objects in the search display) has little effect on response time because individual processing of each item is not required. *Serial* search occurs when the individual items within an array need to be searched one-by-one (or small group by small group) because the target item does not immediately pop out at the observer. Serial searching is needed when the target is only separable from the distractors by a conjunction of multiple features, such that it shares some features with the distractors (e.g., a target red vertical line amongst red horizontal lines and green vertical lines). In these cases, response time increases as the number of items in an array increases because more items need to be searched successively.

This strictly dichotomous view of serial vs. parallel search is no longer considered an accurate characterization of search processes (e.g., Townsend, 1990; Wolfe, 1998), but it continues to offer a useful framework for understanding the variation in processing between simple and complex searches. Because parallel and serial searches are thought to rely on different cognitive processes, it is typically necessary to differentiate between the two when examining an effect, as many conditions may only modulate performance for one of these two types of search.

Feature-integration theory is useful in understanding a simple two-stage concept of the preattentive and focused stages of search, but preattentive processing is more complex than Treisman's original model captures (Wolfe, 1998; Wolfe & Horowitz, 2004). Wolfe's "Guided Search" theory (2007) has a similar, but less linear, model of the stages involved in search. In Guided Search, the basic features serve as guiding attributes to direct the deployment of attention. Both basic sensory processes and selective attention are used in tandem, as basic perception identifies relevant features and guides the observer's attention appropriately. The many versions of the Guided Search model (Wolfe, Cave, & Franzel, 1989; Wolfe, 1994; Wolfe & Gancarz, 1996; Wolfe, 2007) offer a more comprehensive understanding of visual search. These theories of the basic mechanisms of search are important for understanding the underlying processes of visual cognition and allow for analysis of how more complex searches occur.

Bridging the gap between the lab and the field

Historical studies and key cognitive theories of search have built a solid framework for further exploration of exactly what guides visual search performance. A recent trend has been to build upon this framework to approximate critical differences between lab and field searches. In doing so, researchers have purposefully deviated from standard parameters employed in typical lab-based search tasks to introduce factors usually found in field-based searches. For instance, in a typical visual search task in the lab, only one target is possible on any given trial, and half of the trials have a target present. Such standards offer ideal experimental control and statistical power; however, they are not necessarily representative of the nature of field searches where more than one target may be present, and targets may be infrequent. Efforts to better approximate conditions in the field have begun including multiple targets and target categories and decreasing the relative frequency of the targets (e.g., in radiology, Berbaum et al., 1998; Samuel, Kundel, Nodine, & Toto, 1995; Franken et al., 1994; in cognitive psychology, Meneer et al., 2007; Wolfe, Horowitz, & Kenner, 2005; Fleck & Mitroff, 2007; Fleck, Samei, & Mitroff, 2010).

In addition to modifying lab-based search tasks to more directly approximate field conditions, considerable effort has also been dedicated towards advancing technological aids for field searches. Consider, for example, the nature of airport baggage screening; searching X-rays in airports is particularly difficult because of the wide range of potential targets, variability of distractor items, clutter, and potential for purposefully hidden or obscured objects in the search array. When presented with this difficult, but critical, search scenario, it is important to pursue all available means by which to improve performance in both the technology and in the searchers themselves. Key insights have

already been offered in terms of how technological advances may help or hurt the human operator by examining interactions between human factors and technology changes (e.g., Bolfing, Halbherr, & Schwaninger, 2008; Schwaninger, 2006a, 2006b, 2006c; Schwaninger & Hofer, 2004; Schwaninger & Wales, 2009; von Bastian, Schwaninger, & Michel, 2008; Wiegmann, McCarley, Kramer, & Wickens, 2006).

Technological advances can improve field searches along several fronts, but search accuracy still relies on the performance of individual X-ray operators. As such, it is important to study the searchers themselves to find additional ways to increase accuracy. Several research projects have brought the lab and the field together to address this by assessing factors that may both positively and negatively affect search performance (e.g., McCarley & Steelman, 2006; Mitroff & Hariri, 2010; Neider, Boot, & Kramer, 2010; Schwaninger, 2003a; 2003b; Schwaninger, Hardmeier, & Hofer, 2005). Contextual and situational factors potentially present during field searches, such as motivation and anxiety, can impact search processes and performance. Some recent work has examined the effects of motivated and anxious conditions on a variety of cognitive processes (e.g., declarative memory, Murty, LaBar, Hamilton, & Adcock, 2011), but few studies have investigated the interplay of these factors with visual search specifically. Given the numerous cognitive mechanisms underlying successful search and the complexity of many searches in the field, career searchers may be significantly influenced by situational factors that may induce anxiety or increase motivation. Thus, research has begun to explore the impact of context on performance (e.g., Cain, Dunsmoor, LaBar, & Mitroff, 2011) and has determined that such factors can significantly influence search accuracy.

Finally, because of the differences in experience between undergraduates in the lab and career searchers, another technique that has been instrumental in bridging the lab and the field is to test search experts (e.g., radiologists and airport security officers) in a laboratory setting. By controlling for many of the other differences between lab and field, directly comparing the performance between inexperienced searchers and experts on the same task in the same environment allows for the assessment of the effects of expertise on search (e.g., Clark, Samei, Baker, & Mitroff, 2011; Mitroff, et al., 2012). Preliminary results of such studies demonstrate an increase in accuracy with expertise and suggest that the experts employ different strategies.

Many open questions stand in the way of directly translating findings from the lab to the field, and vice versa, but there are good reasons to be hopeful. The current state of visual search research suggests that it is possible to successfully bring a result from one realm to the other, and the current goal is to make this process more and more robust. In Sections 2-5, we discuss four hurdles that present potential problems and strategies for how to overcome them. Specifically, we discuss target prevalence, target number and target category, motivation and anxiety, and level of experience.

2. Target prevalence

When Tyler, our hypothetical undergraduate, begrudgingly sat through a visual search experiment in a dark room, as described in the beginning of this chapter, a target was present on half of the trials he viewed. He did not find every target, but the frequency of targets kept him alert. He was not exerting a significant amount of effort, but the fact

that he was able to find a target so frequently may have helped to keep him on task. If a much smaller percentage of the trials had contained targets, perhaps Tyler would have been more likely to miss those targets.

Olivia, our hypothetical X-ray operator, has rarely encountered actual harmful items in the bags she inspects. Threatening items are, in fact, so rare that the Transportation Security Administration (TSA) has devised a method of inserting images of harmful items onto the viewing screens, superimposed over real luggage. These images, called Threat Image Projections (TIP) (Schwaninger, 2006), are designed to appear as real, dangerous contraband. When the X-ray operator sees a threatening item, he or she pushes the appropriate button on the console. If the item was a TIP image, rather than a legitimately harmful item, the screener is provided with immediate feedback (to avoid detaining an innocent passenger). TIP images are used to keep X-ray operators like Olivia alert, to maintain an index of performance, and to counteract potentially negative effects of the actual low prevalence of targets.

The difference in target prevalence between Tyler's and Olivia's searches highlights a potentially major hurdle: lab-based searches typically have a target present on half the trials, and field-based searches rarely have a target present so often. Lab-based searches use 50% target prevalence levels to maximize statistical power. However, as Olivia experiences, many visual searches conducted in the field do not have this nice balance of 50% target-present and 50% target-absent displays. A termite inspector seldom finds pests on routine inspection; a lifeguard, thankfully, rarely deals with a drowning swimmer; border patrol agents do not routinely see individuals trying to illegally cross a border; and

pilots almost never detect a physical defect on their routine pre-flight structural inspections.

The same case follows for searches in radiology and airport security screening; there is not an abnormality in half of the X-ray images viewed by a radiologist, and there is not a dangerous item in half of the bags viewed by a TSA X-ray operator. The numbers are difficult to calculate for airport security screening, but the rate of truly hazardous items is well below 1% (e.g., Rubenstein, 2001). The prevalence is a bit easier to determine in radiological screening; it is estimated that only around 0.3% of routine mammograms contain an abnormality (Gur et al., 2004). These values deviate substantially from the typical 50% used in the lab, and a critical question is whether the factor of *target prevalence* actually has a functional role in visual search performance. While laboratory search performance is usually quite good, an estimated 30% of malignancies are missed in radiological exams (e.g., Berlin, 1994; Kundel, 1989; Renfrew, Franken, Berbaum, Weigelt, & Abu-Yousef, 1992). Might target prevalence factors account for some of this disturbingly high miss rate? This question is critically important, as failure to identify targets in rare-target search could be potentially disastrous.

Vigilance tasks bear a strong relationship to rare-target search, as they typically consist of a monitoring task in which events occur at rare and unknown intervals, in contrast to a visual search study in which each trial demands a separate response of absent or present. Early studies with vigilance tasks found that performance declines over time while performing a monotonous task (e.g., Mackworth, 1950; Parasuraman & Davies, 1976; Davies, Shackleton, & Parasuraman, 1983). Because rare-target visual searches

resemble vigilance tasks in the monotonous response of “no target,” it is a reasonable assumption that visual search accuracy could decline over the course of time; for example, as screeners repeatedly determine that X-rays do not contain tumors.

Radiological examinations of target prevalence effects have found conflicting results. One study varied target prevalence from 20–60% and found a much higher accuracy rate in higher prevalence conditions (Egglin & Feinstein,1996). However, another study found no difference in performance related to prevalence rates varying from 2–20% (Gur et al., 2003); this lower prevalence rate better maps onto the actual rates of screenings and 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.

Given the complexities of the radiological environment, it not easy to directly assess the role of prevalence with radiologists as the participants and radiographs as the search arrays (e.g., Gur et al., 2003). Likewise, it is not easy to test such questions with X-ray operators and luggage X-rays. Prevalence is, however, possible to address in the lab with inexperienced searchers. Using simplified displays and untrained participants, Wolfe, Horowitz, and Kenner (2005) found a robust prevalence effect. Participants searched arrays of line-drawn objects and were to find “tools” amongst distractor shapes drawn from other categories. Each participant completed searches in which a target “tool” could appear on 50% of the trials (high prevalence), 10% of the trials (low prevalence) and 1% of the trials (very low prevalence). Visual search accuracy significantly declined as the

target prevalence decreased, suggesting that target prevalence, *per se*, may have affected performance (Wolfe et al., 2005). In the 1% target prevalence condition (wherein participants searched 2000 individual trials with only 20 actually containing a tool), participants missed 30% of the targets. While it may just be coincidental, it is nonetheless striking that this number mirrors the probable miss rate from radiology (e.g., Gur et al., 2004). For additional discussion of this point, see Wolfe (2012).

In a typical vigilance task, participants slow down over the course of the experiment (e.g., Buck, 1966). However, the Wolfe et al. (2005) participants were found to speed up over the course of the 2000 rare target trials. It is proposed that as participants repeatedly and continuously correctly reject most target-absent trials, the time taken to reject decreases dramatically. In effect, participants may become so accustomed to saying that no target is present that they stop performing a sufficient search to actually find a target, thus causing a high miss rate on the few target-present trials.

The Wolfe et al. (2005) finding of a target-prevalence effect with simple displays and inexperienced searchers has the potential to be highly relevant to visual searches in the field. If this effect has been properly modeled in the lab, then manipulations can be tested that might improve accuracy (e.g., motivation; Navalpakkam, Koch, & Perona, 2009) and more precise methods can be used to better assess *why* misses occur (e.g., eye tracking; Rich et al., 2008). However, before this lab-to-field link can directly inform visual searches conducted in the field, it is critical to ensure that the link is valid. Does the underlying mechanism of the prevalence effect found in the lab match those of possible prevalence effects in the field?

Follow-up studies have raised concerns about whether the initial prevalence effect found in the lab sufficiently matches prevalence effects in the field (Fleck & Mitroff, 2007; Li, Cao, Xiao, Li, 2011; Li, Chan, Cheung, & Yan, 2012; Madden, Mitroff, Shepler, Fleck, Costello, & Voss, under revision). Fleck and Mitroff (2007) and Li et al., (2011) suggest that the prevalence effect found in the lab may be an error of response execution rather than a perceptual or identification error. When Fleck and Mitroff (2007) offered participants an option to "correct" their responses on a previous trial, this alone removed a previously found prevalence effect. At least in these studies, participants were able to correct such errors, indicating that they were not actually "missing" the targets perceptually; they were simply responding quickly out of habit. In effect, participants fell victim to a classic "oops" problem – they were quickly responding "no" trial after trial, until suddenly they hit the "no" key when, in fact, they had not intended to do so. Such a physical perseveration or inhibition problem is not likely to underlie a prevalence effect in the field. Fast-paced responding is not a common aspect of radiology or baggage screening, and such searches offer the option to correct mistakes.

The results of Fleck & Mitroff (2007) suggest that lab-based visual searches with simple stimuli may not be able to adequately translate to the complex searches conducted in the field, given that the purported mechanism (a response-based error) is not a part of most field searches. An additional study, however, offered a different conclusion. Support for a prevalence effect was found in a study that employed realistic X-ray luggage images (Van Wert, Horowitz, & Wolfe, 2009), even when participants were offered the option to correct their responses. This suggests that a prevalence effect can be observed in the lab,

with the option to correct, as long as the stimuli are sufficiently complex. Moreover, this suggests that prevalence effects are not driven entirely by response-based errors.

A potential problem remains, however, with extrapolating prevalence effects from the lab to the field. While Van Wert et al. (2009) clearly involved more complexity than Fleck & Mitroff (2007) and Wolfe et al. (2005), the locus of the complexity is not as clear. Fleck & Mitroff (2007) and Wolfe et al. (2005) used a set of six possible target "tools" and showed pictures of the targets to the participants before the experiment. Van Wert et al. (2009) used 100 images of knives and 100 images of guns and only showed a small subset to the participants prior to the experiment. While this added desired complexity to the stimuli, it also, unfortunately, added complexity to the participant's task. On 94% of the occasions when participants used the "correction" option in Fleck & Mitroff (2007) to report that they had pressed the wrong response key by accident, they changed misses (responding "no" when a target was present) into hits (responding "yes" when a target is present). In contrast, when the participants in Van Wert et al. (2009) used the correction option in the low prevalence condition, they primarily (81% of uses) changed correct rejections (responding "no" when no target was present) into false alarms (responding "yes" when no target was present). This suggests that the participants in Van Wert et al. (2009) did not understand what was and was not a target and did not have a sufficient grasp of their task. Ultimately, it is not clear what this means for relating prevalence effects from the lab to the field.

While the effects of a correction option on rare-target search performance remain debatable, additional studies have suggested alternative mechanistic accounts of the

prevalence effect that suggest viable connections between lab findings and the field. Further work by Wolfe and Van Wert (2010) demonstrated that not only did searchers' decision criteria shift toward increasing misses at low prevalence, the reverse criterion shift also occurred with very high target prevalence leading to an increase in false alarms. They also found that target prevalence not only influences the criterion shift, but also the decision of when to stop searching in target-absent trials. Another nuanced study examined the prevalence effect in older adults, who typically exert greater top-down attentional control and more cautious approaches when completing search tasks (e.g., Madden, 2007). The older adults not only exhibited less severe prevalence effects but also benefited even more greatly from the ability to correct responses (Madden et al., under revision). The prevalence effect was also found to vary with the number of response alternatives, as the effect was eliminated in a four-alternative forced-choice task but remained intact in the standard two-alternative forced-choice task (Rich et al., 2008). Finally, Lau and Huang (2010) varied instructions given to participants regarding whether there were a high or low number of targets and found that this sort of instruction did not affect performance but that the prevalence effect was driven by the actual distribution of the targets encountered. Furthermore, participants showed the prevalence effect in conditions with a consistent prevalence level, but the miss rate did not increase when the prevalence level varied throughout a block of trials.

These conflicting results highlight the complexities of interpreting visual search performance data. Participants fall victim to the prevalence effect in some cases but not in others. The fact that the prevalence effect differs between younger and older adults

(Madden et al., under revision) demonstrates that inherent differences between participants affects search performance and provides a note of caution when trying to translate results from undergraduate searchers in the lab to older, professional searchers in the field. When these findings are taken together, it becomes apparent why interpreting visual search data is almost never straightforward. Both the characteristics of the experiment and the characteristics of the participants can dramatically alter performance results, such that isolating an effect to one specific cause is often impossible. Using a rare-target search task is a far more comparable means by which to consider search performance in the field, but it must be done in an informed way while taking all nuances discussed here into consideration.

3. Number of targets and number of target categories

Tyler, our hypothetical undergraduate, is aware there is never going to be more than one target-T shape within any display he views. His task is rather simple – he searches for a single target of a single category, and once he finds the target, his search is complete. Tyler does not need to concern himself with additional targets nor additional target types. He is not required to maintain two (or more) separate templates in memory while searching, and once he finds a target, he knows no further searching is required.

Olivia, our hypothetical X-ray operator, is tasked with searching for multiple kinds of items at all times. Not only does she need to identify guns, knives, and bombs, but she also needs to search for other items such as laptops, shoes, and liquids within each bag she examines. Furthermore, when she finds one target in a bag, she cannot terminate her

search after the identification of this single target because there is no limit to the number of harmful items potentially present.

The possibility of multiple targets highlights another critical question: does it matter if someone is searching for more than one thing at the same time? The majority of lab-based visual search tasks present participants with well-defined stimuli and ask them to search arrays that contain either zero or one target. However, searches in the field can often contain more than one target type (e.g., either a tumor *or* a broken bone in a single radiograph) and/or more than one target (e.g., a tumor *and* a broken bone in a single radiograph). These types of searches, which we will refer to as *multiple-category* and *multiple-target* search, respectively, are rarely employed in the lab but are frequently present in the field. Is it possible to generalize from lab-based single-target research to multiple-category and multiple-target field-based searches? What is the cognitive cost of having to maintain in memory more than one target type? Is search performance worse if there may be multiple targets in the same display? We discuss these questions in this section and explore how they may present hurdles for translating research between the lab and the field.

Multiple-category visual search.

An X-ray operator is tasked with finding dangerous items and must simultaneously search for guns, knives, bombs, water bottles, and several other potentially dangerous items. Multiple-category visual search has a long history in cognitive psychology (e.g., Kaplan & Carvellas 1965; Krueger & Shapiro, 1980; Menneer et al., 2004; 2007; 2008;

2009; Neisser, Novick, & Lazar, 1963; Vreven & Blough 1998), and several conclusions have emerged. It is clear that there is a negative impact of having to hold more than one potential target in memory (e.g., Gould & Carn, 1973). When varying the number of possible target categories, larger numbers of target categories led to steeper search slopes (Kaplan & Cavellas, 1965) and slower searches overall (Metlay, Sokoloff, & Kaplan, 1970).

Kyle Cave and his colleagues have convincingly shown that multiple-category search has a detrimental effect in terms of both visual search speed and accuracy (e.g., Meneer, Barrett, Phillips, Donnelly, & Cave, 2007; Menneer, Cave, & Donnelly, 2009; Godwin, Menneer, Cave, & Donnelly, 2010). In one study, different groups of participants searched for either one or two colors, one or two shapes, or one or two line orientations (Meneer et al., 2007). Search times were slower, and miss rates were drastically higher in the dual-category trials. In a study that was directly inspired by airport baggage screening, participants had to search X-ray images for either weapons *or* bombs in separate searches or weapons *and* bombs in the same search (Godwin et al., 2010). They found that there were dual-category search costs in both accuracy and response time and that low-prevalence targets were missed more often than high-prevalence targets, but these factors appear to be additive. This suggests that searchers in the field who are searching for rare targets in many categories may be subject to many sources of miss errors.

The aforementioned work highlights two important points for the current discussion. First, multiple-category search is considerably slower and more error-prone than single-target visual search. Second, lab-based research can inform—and be informed by—visual searches in the field. Experiments using both simplified stimuli (e.g., Menneer

et al., 2007) and X-ray baggage images (e.g., Godwin et al., 2010), have revealed ways in which multiple-category search demands impact performance and have offered suggestions for how to improve real-world searches. For example, Menneer et al. (2007) suggests that because simultaneously searching for multiple types of targets (e.g., guns, knives, and bombs) produces costs for both search speed and accuracy, it may be more effective to have multiple, specialized searchers that are focused on a single target type (e.g., only guns or only bombs).

Multiple-target visual search findings from the lab.

Multiple-category visual search requires holding more than one item in memory but does not necessarily involve identifying more than one target within the same array. Once a target has been found in a single-target search, the search can immediately be terminated, but what happens when there are potentially more targets? Does the successful detection of one target make a searcher more likely to notice additional targets? Or, does it act as a distractor and impair further search? These are critically important questions, as many visual searches in the field—where misses can be disastrous—have an unconstrained number of targets.

The nature of multiple-target search has been directly examined in the lab in a few studies (e.g., Cain et al., 2011; in press; Chan and Courtney, 1995; Fleck et al., 2010; Schneider & Shiffran, 1977; Wolfe, Horowitz, & Kenner, 2005). One series of studies examined multiple-target visual search accuracy using an array of measures: useful field of view (Chan & Courtney, 1995), target discriminability (Chan, Courtney, & Ma, 2002), and

the time course of search (Chan & Chan, 2000). In the first of these experiments (Chan & Courtney, 1995), participants were briefly (250 ms) shown a horizontal line of shapes and were to report whether any 'o' shapes were present. The majority of the shapes were 'x's, but there were a variable number of target 'o's. When there were two targets present, participants were more likely to report the target that was presented closer to center and less likely to report the target presented in the periphery than when targets in those same locations were presented as the only target on a trial (Chan & Courtney, 1995). This result implies that multiple-target search effectively reduced searchers' useful field of view, compared to single-target searches. In a version of the task designed to look at the time course of multiple target search, participants again scanned an array of 'x's for '<' and '>' but with both targets present on all trials. The time taken to find a second target was much more variable than that needed for the first target (Chan & Chan, 2000), suggesting that modeling a dual-target search as two, serial single-target searches would not properly reflect actual search behavior. In another accuracy-focused study, participants searched for either two hard-to-spot shapes ('<' and '>') or an easy-to-spot and a hard-to-spot shape ('o' and 'v') among 'x's. The presence of a hard-to-spot target impaired detection of an easy-to-spot target more so than the presence of an easy-to-spot target impaired detection of a hard-to-spot target (Chan, Courtney, & Ma, 2002), reinforcing similar findings from radiology (Berbaum et al., 2001).

The above studies suggest that several factors can impact multiple-target search accuracy, and an additional study has suggested that the top-down knowledge of a multiple-target search can affect search even before the first target is located. Körner and

Gilchrist (2008) compared eye movements between a condition in which there were 0 or 1 targets present and a condition in which there were either 1 or 2 targets present, with participants informed of the conditions. Even before a target was found, participants made more distractor re-fixations in the 1-vs.-2 condition on the trials with just 1 target present than on physically identical 1-target stimuli in the 0-vs.-1 condition. This difference was argued to arise from participants "setting aside" memory for a possible second target before the search began, thus limiting the available memory for which locations have been searched (Körner & Gilchrist, 2008). This finding suggests that, not only do physical aspects of the search array affect performance, but that searchers' expectations about the likely number of targets may also affect the efficiency of their search.

Multiple-target search in radiology: "Satisfaction of Search."

The problems accompanying the presence of more than one target and different types of targets (e.g. pulmonary nodules and fractures) have been well documented in radiology (e.g., Berbaum et al., 1998; Samuel et al., 1995; Franken et al., 1994). A classic pitfall, known as "satisfaction of search" (SOS), occurs when the identification of a second target is less successful after the identification of a first target in the same display. SOS has been a topic of radiological research since the 1960s (Tuddenham, 1962), but radiologists still fall victim to SOS. In fact, 28% of radiological misses have been attributed to SOS errors, which makes this a critically important problem to solve.

Radiography studies have delineated three possible types of errors contributing to SOS (Nodine & Kundel, 1987): *scanning errors* (the search path never encounters the

target area, Berbaum, 1996, 2005; Samuel et al., 1995), *recognition errors* (scanning in the region of a possible target but failing to dwell on the correct area for further inspection, Berbaum, 2000), and *decision-making errors* (fixating and dwelling on a possible target but ultimately failing to identify it as a target, Franken, 1994). To date, evidence has suggested that all three likely contribute to SOS and the latter two explanations differ primarily in the amount of time spent analyzing a potential target. The time required to examine a target stems, in part, from the relatively low spatial frequency of radiographs and radiological targets (e.g., pulmonary nodules), which may require extra analysis to visually parse targets from background noise. By categorizing errors as scanning or decision-making/recognition errors, radiologists have attempted to understand whether SOS arises primarily as a function of a basic perceptual failure to properly scan an image or more of a cognitive failure in determining whether a particular item is indeed a target (Kundel, Nodine, & Carmody, 1978).

Several suggestions have been proposed for the causes of multiple-target search errors, and some possibilities include a *truncated search* (finding one target leads to a non-exhaustive search) and a *perceptual set* (e.g., once a tumor is detected, the searcher engages a "tumor set" where additional tumors are likely to be spotted but other abnormalities, for example a fracture, are less likely to be spotted, Berbaum et al., 2000). However, no clear mechanism has been identified as responsible for the SOS effect. Eye-tracking data within radiological research has indicated that search is not actually terminated early (Samuel et al., 1995) and that participants continue to search after the successful identification of a first target; participants may even fixate on a second target

but fail to identify it as such. This is confirmed and extended by a recent eye-tracking study done with undergraduate searchers and simplified search displays (Cain, Adamo, & Mitroff, 2012). The eye-tracking data suggest that SOS errors are likely due to a combination of scanning, recognition, and decision-making errors.

The stimuli used in radiological search studies, however, are actual radiographs, which are highly complex and greatly variable. Furthermore, participants are radiologists who have extensive training and experience with the experimental tasks. In contrast, many lab-based visual search tasks do not involve a heavy decision-making component because the targets and distractors are easily distinguishable from one another (in part so that untrained participants can perform reasonably). Recent work has begun to explore the SOS effect outside of radiology using simplified stimuli and undergraduate participants rather than radiographs and radiologists (Fleck, Samei, & Mitroff, 2010) in order to generalize SOS to outside of the medical world.

In seeking to establish the scope of SOS errors in nonmedical searches, Fleck et al. (2010) aimed to understand the cognitive processes broadly involved in multiple-target search. In a series of experiments, basic 'T' and 'L' shapes were presented at varying degrees of visibility against a cloudy background. The cloudy background and the more- and less-salient targets and distractors served to approximate the noise typically present in radiographs. See Figure 1. Each trial contained 0, 1, or 2 targets, and participants were to make a localization mouseclick on each target found and then clicked a 'DONE' button at the bottom of the screen to terminate their searches.

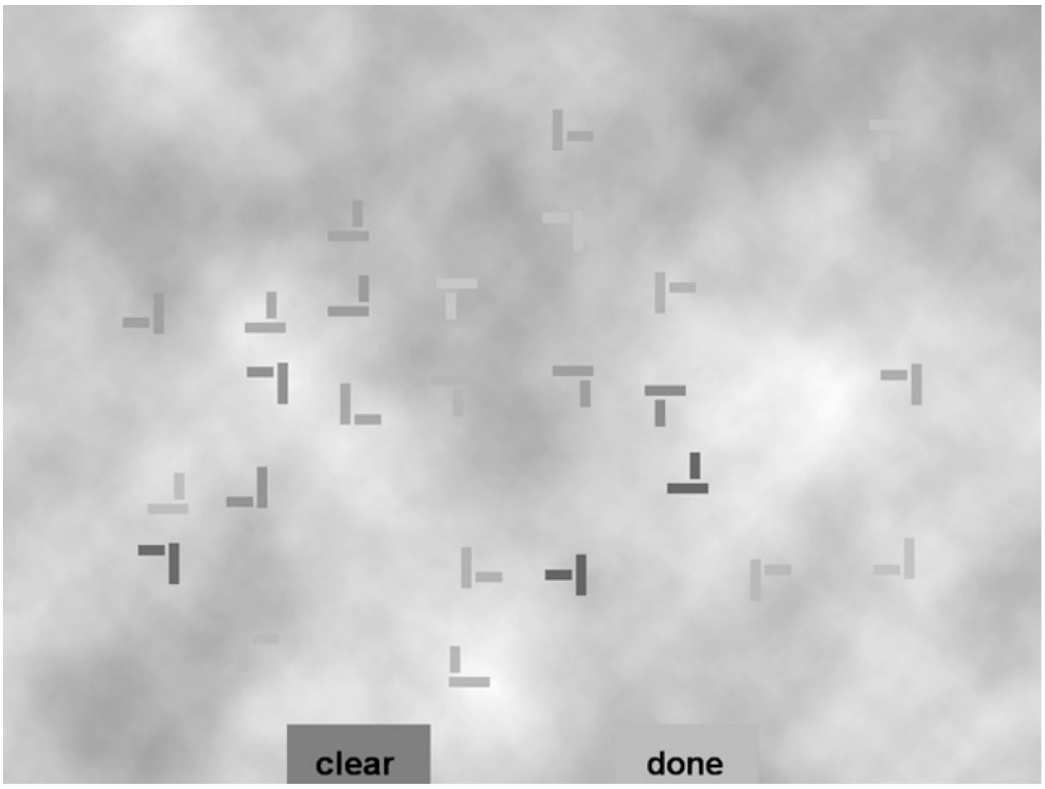

Figure 1: Sample trial: Find the "perfect" T shapes.

In previous studies, the SOS effect was typically observed when radiologists were less likely to identify a low-salience target when it was in the presence of a high-salience target than when the same low-salience target was the only target present in the array. Thus, the dual-target trials in this paradigm contained both a low- and high-salience target, and the SOS effect was calculated as the difference between a participant's accuracy in identifying low-salience targets in single-target trials and his or her accuracy in identifying low-salience targets in dual-target trials, provided the high-salience target had been successfully identified in the same trial. The SOS effect was found to be sensitive to both target prevalence and time pressure, as it was exacerbated when high-salience

targets were three times as likely as low-salience targets and when participants had a time limit of 15 seconds per trial (Fleck et al., 2010).

These findings demonstrate generalized SOS errors in nonmedical searches; the inclusion of the possibility of multiple targets allows this paradigm to be applied to the field, as it is clear that search processes become infinitely more complex when multiple targets may be present. In order to gain a full understanding of search processes employed in the field, multiple-target searches should be used when exploring other issues related to lab-field differences.

4. Motivation and anxiety

Tyler, our hypothetical undergraduate student, is simply not very concerned with his performance on his computer-based visual search experiment. He has little reason to care if he finds every target; regardless of how he performs, he will receive the participation credit for his Introductory Psychology class. There are also no consequences for poor accuracy. In addition to a lack of motivation, he likely has little to no anxiety about how well he does on this task since this really is just some "meaningless" experiment to him.

On the other hand, Olivia, our hypothetical X-ray operator, is strongly motivated to perform well since her job security is, at least partially, based on good performance and accurate searching. She takes great pride in her search abilities, as she has been with the TSA for quite some time and is a seasoned X-ray operator. More importantly, she is well aware of the consequences that might accompany her failure to identify harmful items in

her search, and this keeps her motivated. Unfortunately, Olivia finds that she is often quite anxious while at work since passengers are always around her, and they are usually visibly (and often verbally) annoyed and hurried.

Difference in context: Motivation.

In the above scenarios, it is clear that Tyler and Olivia are faced with wildly different motivational contexts. Tyler's performance has no impact on his life, and there is little reason, beyond personal pride, to perform well. Olivia's performance can impact her livelihood (e.g., whether she has a job in the future) and others' lives (e.g., whether they are boarding a plane along with a bomb). Searches in the field are often linked to high-stakes outcomes; a radiologist or X-ray operator could save lives by identifying harmful targets in X-rays. Does performing a life-critical search cause individuals to be more motivated than when completing a lab-based task with no tangible consequences? The primary issue addressed in this section is whether differing levels of motivation affect visual search performance, and if so, how. If higher levels of motivation result in higher levels of performance, then how comparable are unmotivated, inexperienced searchers to highly motivated career searchers? This is a third fundamental hurdle for translating findings from the lab to the field.

Visual search tasks conducted in the lab often reveal a great deal of variability in performance in undergraduate participants. While some of this variability may be tied to differences in underlying search ability, some variability may also result from differences in motivation: Some participants may be intrinsically motivated to perform well regardless

of a tangible outcome while others may not be motivated at all. For instance, more conscientious participants are likely to exert greater care and effort when performing the task, even though their levels of performance have no external consequences for them.

Unfortunately, it is not feasible to directly motivate laboratory participants in the same way career searchers are motivated. Participants cannot possibly believe that people's lives are in their hands nor that their careers depend on their performance in a computer-based experimental task. One reasonable approximation of motivation, however, is performance-based monetary reward. The prospect of receiving money for good performance provides an effective global incentive that, for most people, will increase their interest and effort (e.g., Camerer & Hogarth, 1999).

In typical lab-based visual search experiments, monetary reward has been used to examine the impacts of motivation on attentional selection (e.g., Libera & Chelazzi, 2006; Kiss, Driver, & Eimer, 2009), priming (e.g., Hickey & Theeuwes, 2008; Kristjansson, Sigurjonsdottir, & Driver, 2010) and attentional capture (e.g., Anderson, Lauren, & Yantis, 2011). These studies have presented clear evidence that monetary rewards can improve performance; however, they have primarily focused on changes in the speed of attentional deployment. While this is a critical component of visual search performance, the majority of field-based searches place a larger emphasis on accuracy than on speed. Two recent experiments have employed monetary incentives with a focus on visual search accuracy: one with rare-target visual search (Navalpakkam, Koch, & Perona, 2009) and one with multiple-target visual search (Clark, Cain, Adcock, & Mitroff, 2011).

Navalpakkam et al. (2009) investigated whether the prevalence effect (that targets are missed more when they occur rarely than when they occur frequently) could be overcome when participants were sufficiently motivated. Participants searched for a target object in a cluttered scene, with the target prevalence (2%, 10%, and 50%) varied across blocks. A typical pattern emerged, with impairments in accuracy at low target-prevalence (e.g., Wolfe et al., 2005). However, when participants were motivated with a monetary incentive, the prevalence effect decreased significantly, restoring detection rates to near optimal levels. It was argued that fatigue, carelessness, and lack of vigilance were not responsible for the prevalence effect, but instead, the prevalence effect was caused by a shifted decision criterion, which could be modified through proper reward (Navalpakkam et al., 2009).

In Section 3, we described the pitfalls of multiple-target visual search: searchers are less likely to find a target if they have already found another target in the same display (a phenomenon termed "satisfaction of search," SOS). This is a potentially dangerous problem that has been consistently observed in both lab-based and field-based visual searches (e.g., see Berbaum et al., 2010; Fleck et al., 2010). A recent series of experiments (Clark, et al., 2011) has explored whether SOS errors can be alleviated with the incentive of monetary reward. Can certain motivational frameworks lead to performance differences, and do such differences provide information on how to better structure work conditions for career searchers?

Clark et al. (2011) employed a multiple-target search tasks that mirrored a paradigm that has previously found robust SOS errors (Fleck et al., 2010, Experiment 3,

described in more detail in Section 3) but manipulated the participants motivation by including a monetary incentive. Participants competed against 9 other participants, and the "best" performer was awarded an additional $50. By simply adding this motivation of a performance-based reward, accuracy improved and the SOS effect was effectively eliminated.

These results raise a curious concern for the translatability of research considering that motivated, inexperienced participants show a decreased SOS effect, yet career searchers such as radiologists, presumably operating while motivated, still exhibit the SOS effect. Clearly, the incentive for the inexperienced participants differs greatly from the incentives for career searchers, but one could argue that a relatively small monetary incentive is qualitatively less than the incentive for career searchers—the chance at winning an extra $50 at some point in the next week or two is seemingly less motivating than keeping a job and preventing fatalities. How can this be reconciled? Perhaps despite the immense focus on accuracy for career searchers, the monotony of their daily routines interferes with their motivation. It is possible that the undergraduate searchers could actually be more motivated than the career searchers because they are completing a task for only an hour-long period, over which it is relatively easy to maintain a high level of motivation. Career searchers may not be not equally motivated at every hour throughout their workdays (or weeks, or months, etc.), and the SOS effect is observed may result from an inability to maintain consistently high levels of motivation.

Differences in context: Anxiety.

The dire consequences of missing a target in field searches could be potentially motivational but could also induce anxiety. Anxiety—the displeasurable psychological experience of worry or concern—is difficult to replicate in the lab, but it may be an element in many field searches. Beyond the general anxiety of knowing that missed targets could have life-threatening consequences, there is also more acute anxiety that can occur when searchers anticipate tangible stressors, such as a visit from a supervisor or a large workload. These states of heightened anxiety can be detrimental to accuracy, and anxiety has been linked to a decline in cognitive performance across species (e.g., in mice, Ohl, Roedel, Binder, & Holsboer, 2003; in humans, Eysenck, Derakshan, Santos, & Calvo, 2007).

As discussed earlier in this section, the motivation to earn rewards can significantly improve performance (e.g., Callan & Schweighofer, 2008; Murayama & Kuhbandner, 2011), but motivation to avoid punishments can increase anxiety and substantially diminish performance (Davis & Whalen, 2001; Lang & Bradley, 2009). Recent work has examined the differing effects of *approach* and *avoidance* motivation (earning rewards and avoiding punishments, respectively) on declarative memory (Murty, LaBar, Hamilton, & Adcock, 2011). While approach motivation enhanced memory performance, avoidance motivation hindered performance, and this effect was especially amplified in participants who showed high levels of arousal. In the lab, the experience of *anticipatory anxiety* can be induced using a "threat of shock" paradigm, in which electrical shocks are administered at unpredictable intervals, unrelated to performance (e.g., Grillon, Baas, Lissek, Smith, & Milstein, 2004; Rhudy & Meagher, 2000). A recent study has found this

type of anticipatory anxiety to be specifically detrimental to multiple-target visual search performance (Cain et al., 2011).

Cain et al. (2011) used a variant of a standard multiple-target search paradigm (Fleck, et al., 2010, experiment 5) in which an SOS effect was *not* expected. When participants were anticipating a neutral event (an innocuous tone), they did not show SOS (which replicates the previous instantiation of these particular experimental parameters). However, when those same participants were anticipating a negative event (an electrical shock) they produced SOS errors. Interestingly, the participants did not show a difference on single-target performance between the non-anxious and anxious blocks of trials—the SOS effect was due solely to poorer second-target identification (Cain, Dunsmoor, LaBar, & Mitroff, 2011). Moreover, this effect was modulated by the level of anxiety that participants were experiencing at the start of the experiment. Less anxious participants showed high levels of SOS when anticipating a shock but no SOS in the control condition, while more anxious participants showed mild SOS throughout the entire experiment, regardless of condition. These results suggest that both acute and generalized anxiety could negatively affect search performance in the field by inducing SOS errors. Thus, efforts should be made to shield professional searchers in the field from anticipatory anxiety in order to improve target identification in multiple-target displays. This sort of anxiety potentially poses an extra risk of misses in searchers with post-traumatic stress disorders or clinical anxiety disorders given that these individuals have been shown to be more likely to generalize specific causes of anxiety to the environment itself (e.g., Fanselow, 1980; Grillon, Morgan, Davis, & Southwick, 1998).

Just as the prevalence effect can be overcome with the proper motivation, multiple-target search appears to be influenced by contextual conditions. Performance on single-target searches in both motivated (Clark, et al., in preparation) and anxious (Cain, et al., 2011) conditions were unaffected by context, and influences were seen only on multiple-target conditions. The complex mechanisms responsible for the SOS effect may simply be more sensitive to contextual influences, and the motivation and anxiety inherent in career searches may work both for and against performance.

Though SOS can be eliminated in the laboratory via monetary incentive, SOS remains a problem in the field. This could be attributable to the monotony of the daily grind detracting from the value of motivation, but the anxiety associated with career searching could also contribute negatively. While motivation appears to positively affect performance, anxiety may serve as a hindrance, and it is important to take all of these factors into account when evaluating differences in searching between the lab and the field. Furthermore, enhancing motivation while decreasing anxious circumstances may be the best combination of contexts for optimal search performance.

5. Level of experience

Tyler, our hypothetical undergraduate, searches for items in his everyday life – the books for his classes, the keys to his dorm room, and his cell phone. He rarely, if ever, dedicates any sort of mental effort toward improving the efficiency of these searches since they are mundane and generally completed successfully. Tyler has also never been trained on how to conduct visual searches to increase accuracy, nor does he regularly spend

hours at a time conducting visual searches (e.g., he usually finds his keys within a minute or two at the most).

Olivia, our hypothetical X-ray operator, has undergone extensive training in order to improve her search skills. Additionally, she spends hours every workday actively conducting visual searches as she scans bag after bag. She has years of experience in search that have allowed her the opportunity to increase her ability to scan X-ray images for harmful items, allocate her visual attention more effectively, and utilize any superior strategies she may have developed.

The above vignettes about Tyler and Olivia highlight the last major hurdle we discuss in this chapter—experience. Many career searchers have years of training and experience on specific search tasks, and it is important to understand how this might influence their abilities. It is not clear exactly how career searchers' levels of experience may affect their performance, both on their typical job-related searches and on search tasks more generally. How might search expertise on the job translate to search performance on standard lab-based search tasks? Which conclusions drawn from inexperienced undergraduates, without extensive training, are applicable to the field?

Trained professionals are often better at visual searches related to their jobs than are novices (e.g., farmers improve their ability to sort chickens by sex with experience, Biederman & Shiffrar, 1987; wine connoisseurs learn to discriminate between fine wines, Bende & Nordin, 1997; bank tellers are better than the general public at detecting counterfeit currency, Klein, Gadbois, & Christie, 2004; and chess players are better able to see patterns of moves on a chessboard, Chase & Simon, 1973). This apparent benefit of

experience leads to two key questions: What are the bases for these expertise differences? And how can the differences be accounted for when assessing the performance of inexperienced searchers in an attempt to translate from the lab to the field? Observing how expertise may alter both trained task performance specifically, and visual/cognitive abilities more generally, has the potential to inform questions about visual search as well as the general malleability of cognitive abilities.

Perceptual training in the lab.

In most cognitive psychology studies, a participant (like our hypothetical undergraduate, Tyler) arrives in the lab, runs through a minute or so of practice, and then completes an hour-long study. They are then dismissed and may never think about the task again. The experimental results provide a useful assessment of performance but do not allow for an investigation of learning. One class of experiments, however, is focused primarily on learning effects. In *perceptual learning* experiments, a research participant may make several visits to the lab and undergo thousands of trials of the same specific task so that they ultimately receive extensive training.

Research in perceptual learning has shown that it is possible for very basic visual abilities to change with experience. If a participant is asked to make a difficult visual discrimination over many trials—often spread out over days—his or her threshold for discrimination will decrease dramatically (Westheimer & McKee, 1978); these changes are often attributed to plasticity in primary visual cortex (V1). There has been evidence for a host of sensory and perceptual improvements in which basic feature discrimination

improves with extensive practice of a task. Participants improve in discriminating the orientation of a line (Ramachandran & Braddick, 1973; Fiorentini & Berardi, 1981; Matthews & Welch, 1997), identifying the direction of motion (Ball & Sekuler, 1982; 1987), and show increased vernier acuity (Westheimer & McKee, 1978; Saarinen & Levi, 1995; Beard, Levi, & Reich, 1995). In all of these cases and in others (e.g., Vogels & Orban, 1985; Karni & Sagi, 1991; 1993; Poggio, Fahle, & Edelman, 1992; Fahle & Edelman, 1993), learning is specific to the stimulus on which the participant was trained. In fact, a hallmark aspect of perceptual learning is that the training effects appear to be quite specific. Because the learning is believed to take place at such a basic perceptual level, improvement is only seen when examining performance on the exact trained stimulus. If, after training, participants showed an overall improvement in a task, beyond that of the trained stimulus (e.g., he/she was trained to identify rightward motion, but also improved in identification of leftward motion), more generalized training would be said to have occurred.

Generalized training.

Perceptual learning studies in the lab have produced highly specific training effects, but this does not directly inform generalized learning effects. Given the uncontrolled and variable nature of field-based searches, expertise gained through career searching likely produces more generalized benefits. In airport security screening, for example, X-ray operators never search two entirely identical suitcases, so they cannot rely on simple sensory-level template matching to successfully identify targets. The experience gained

through their daily training relies on improvements that *can* transfer from bag to bag. Furthermore, visual search, even its most simplistic, laboratory form relies on the integration of both sensory perception and strategic attentional allocation.

One of the few perceptual learning studies to demonstrate generalized learning used a visual search task (Sireteanu & Rettenbach, 1995). Training accumulated over the course of the experiment resulted in improved search efficiency, even on untrained stimulus sets. Perhaps because of the complex attentional processes required for effective visual search, the learning occurred in a less specialized manner; visual search involves cognitive processes more complex than basic sensory discrimination, so the improvement likely occurred at a level that can generalize beyond the perception of one specific stimulus. For example, search efficiency may improve via changes in strategies—participants may learn to better distribute their attention, disregard irrelevant cues, or react quickly to relevant ones.

Because search relies on strategies and attentional processes, it is impossible to isolate learning for visual search tasks to the sorts of low-level feature discrimination improvements seen in classic perceptual learning tasks. As such, "perceptual learning" has recently undergone a redefinition, which encompasses even strategic and attentional improvements. Using visual search paradigms to study trained improvement allows for a more comprehensive investigation of the many elements of attention that can be improved with training.

Generalized learning via action video game playing.

Basic sensory perceptual learning cannot allow for improvement in skills in the field, where search arrays consistently vary. Certain experiences, however, have been found to elicit improvement in a wide variety of skills and are far more generalized than basic perceptual learning processes. Extensive experience with specific activities can influence perceptual and attentional abilities that generalize beyond those activities, and a host of studies have shown that those who regularly play action video games (usually an average of 6 or more hours per week for at least 6 months) show improved performance on a variety of tasks. Specifically, when compared to those who did not regularly play action video games, avid action video game players respond more rapidly (Castel, Pratt, & Drummond, 2005; Dye, Green, & Bavelier, 2009; Orosy-Filders & Allan, 1989; Yuji, 1996), have improved spatial abilities (Okagaki & Frensch, 1994; Quaiser-Pohl, Geiser, & Lehmann, 2006; Terlecki & Newcombe, 2005), have enhanced temporal abilities (Donohue, Woldorff, & Mitroff, 2010; Green & Bavelier, 2003, 2006b, 2007; West, Stevens, Pun, & Pratt, 2008), can enumerate briefly displayed items more quickly (Green & Bavelier, 2006b), can switch between tasks faster (Cain, Landau, & Shimamura, 2012; Karle, Watter,& Shedden, 2010), and have enhanced eye–hand coordination (Griffith, Voloschin, & Gibb, 1983).

Studies exploring the causal role of video game playing have trained non-gamers on action video games and shown improved performance (e.g., De Lisi & Cammarano, 1996; De Lisi & Wolford, 2002; Dorval & Pepin, 1986; Green & Bavelier, 2003, 2006a, 2006b, 2007; however, see Boot Kramer, Simons, Fabiani, & Gratton, 2008 for lack of training effects; and Nelson & Strachan, 2009 for more nuanced training effects). The issue

of causality explores an important mechanistic explanation of gamers' benefits, but regardless of the causal nature of such benefits, differences between gamers and non-gamers have been reliably demonstrated.

However, there is a mechanistic question regarding these differences; two feasible accounts have both received support and are not mutually exclusive. The basic-sensory hypothesis suggests that action video game exposure trains better "vision" and "attention," honing basic abilities (e.g., Dye et al., 2009; Green & Bavelier, 2006a; 2007; Li et al., 2009; West et al., 2008; Caplovitz & Kastner, 2009). According to this hypothesis, gamers may have an increased capacity to process visual information compared to non-gamers. Alternatively, the improved-strategy hypothesis suggests that video game playing leads to the development of enhanced higher-level abilities such as attentional control (Cain et al., 2012; Chisholm et al., 2010; Hubert-Wallander, Green, & Bavelier, 2010), shifts in attentional allocation, and improved strategy (Clark, Fleck, & Mitroff, 2011) for generalized use across a variety of visually demanding tasks. In line with this account, gamers need not necessarily have an increased information-processing capacity but rather could be better able to use what resources they have to process perceptual information (e.g., Colzato, van Leeuwen, van den Wildenberg, & Hommel, 2010).

Generalized learning via stroboscopic training.

In addition to video-game learning, stroboscopic training has been shown to improve visual cognition abilities (Appelbaum, Schroeder, Cain, & Mitroff, 2011; Appelbaum, Cain, Schroeder, Darling, & Mitroff, under review). Stroboscopic, or

intermittent, vision is the process of presenting an individual with snapshots of the visual environment rather than a continuous visual experience. Training in such a visual environment can alter perceptual-motor abilities (e.g., Bennett, Ashford, Rioja, & Elliott, 2004; Mitroff, Friesen, Bennett, Yoo, & Reichow, under review; Smith & Mitroff, under review), and recent work suggests it can influence visual attention and memory as well. For example, in Appelbaum et al. (2011), participants trained on sports activities (e.g., playing catch) while either wearing transparent eyewear or stroboscopic eyewear that occluded vision at regular intervals. Before and after training, participants completed computer-based tasks without the eyewear. In one task, participants viewed patches of moving dots presented either centrally or peripherally and reported which of two sequentially-presented patched had coherent motion. Those participants who wore stroboscopic eyewear during training showed greater test-retest improvements on motion coherence sensitivity for centrally presented patches than participants who wore transparent eyewear, but no effects were seen for peripherally presented motion (Appelbaum et al., 2011). In another task, a useful field of view experiment, participants were briefly (~90 ms) shown a central letter and a dot in one of 24 peripheral locations. After a masked delay, they were asked to report the location of the dot and whether the central letter was upper or lower case. While the central task was primarily intended as a fixation control, the participants who trained with stroboscopic eyewear showed significant test-retest improvement at accurately reporting the case while the control group did not. No differences were found in peripheral performance for either group. Taken together, these results suggest that stroboscopic training may lead to generalized

perceptual improvements, particularly in the center of the visual field (Appelbaum et al., 2011).

Career training and visual abilities.

Another form potentially generalized training occurs on the job; radiologists, for instance, spend years learning how to properly scan radiographs. Medical searches are among the most commonly studied visual searches in the field and have provided evidence for both specialized and generalized learning. Trained orthodontists are better able to detect subtle facial asymmetries than general dentists or lay people (Kokich, Kiyak, & Shapiro, 1999), suggesting a specific enhancement in their trained skill set. Similarly, radiologists and cytologists are better able to detect abnormalities in medical images than inexperienced searchers, but they do not have enhanced memory for these abnormalities nor are they better at simple scene or object detection, suggesting specific training benefits. However, surgeons who regularly engage in video game playing were found to perform better at assessments of laparoscopic surgery (Rosser et al., 2007), suggesting a generalized benefit.

Radiologists typically have years of experience searching medical radiographs for abnormalities, but research in radiology shows they still fall victim to many of the same types of errors as inexperienced searchers. The studies of radiological visual search described in Section 3 focused on the satisfaction-of-search (SOS) effect using radiologists as participants and real radiographs as test stimuli. Using actual radiologists and the stimuli they normally view to address research questions is entirely sensible, but it limits

the ability to compare performance across different real-world expert populations. Showing an inexperienced searcher a radiograph or an X-ray of a bag may not be the best way to assess abilities since there would be a baseline difficulty that could mask their performance. One way to more directly compare abilities from the lab to the field is to use simplified displays for all participants, and a few recent studies have done so. One study used simple visual search arrays to study SOS in both undergraduate participants and training radiologists (Clark, Samei, Baker, & Mitroff, 2011), and another used simple visual search arrays to compare and contrast undergraduates and working airport baggage screeners (Mitroff et al., 2012).

To compare performance between radiologists and inexperienced searchers, Clark et al. (2011) administered a simplified multiple-target search task (e.g., Fleck et al., 2010) to both trained radiologists and to undergraduate students. A broad analysis of the data indicated that, perhaps surprisingly, radiologists and inexperienced searchers did not differ in overall search accuracy; the percentage of trials that they completed correctly (no misses, no false alarms) was not significantly different. However, the radiologists spent significantly longer per trial than did the undergraduates, and they frequently exceeded a trial time limit (15 seconds), while the undergraduates searchers hardly ever did so. The inexperienced searchers were actively deciding they had finished searching and electing to terminate their searches, while the expert searchers may have just run out of time while attempting to complete a more thorough search. When only assessing performance on trials in which participants indicated they had completed the search before the time limit, radiologists were more accurately able to detect the presence of a second target in a

display, showing a reduced satisfaction of search effect relative to the inexperienced searchers.

Returning to the key hurdle discussed in Section 4, different levels of motivation between searchers in the lab and in the field, it is worth considering whether the above differences between radiologists and inexperienced searchers might stem from radiologists simply caring more about their performance. Perhaps they took longer to respond and were more accurate because they were more motivated to perform well. This is a general concern for any such comparison, and one way to address this is to have the participants perform an additional, orthogonal task that does not tap into the specific skill in question. This was done in this study, with all participants also completing a control task, on which they made judgments about the temporal order of appearing squares. No differences between groups were found on this control task, which helps dampen the motivational concerns.

Experience appears to dramatically impact performance on cognitive tasks, but not in a simple, straightforward manner. Because of the perceptual variability from X-ray to X-ray, it is unlikely that any improvement would stem from enhanced basic sensory abilities. Instead, it appears that improvement in strategy or better attentional allocation may contribute to the improvement that comes with experience. Expert searchers are going about their searching in very different ways from inexperienced searchers and are likely more effective as a result. However, the mechanisms responsible for these differences remain largely unclear.

6. Discussion and Conclusions

The goal of this chapter was to explore ways in which visual search findings can translate between the lab and the field. On one side, a tremendous number of visual search experiments have been conducted in the lab, and the data have served as the basis for intricate and powerful theories of search. On the other side, career searchers conduct visual searches daily and are constantly looking for ways to improve performance. The critical question is whether each side can inform the other. Can cognitive theories and data be used in the field to guide and inform search practice? Likewise, can the nature of field-based searches be analyzed to further refine cognitive theories?

At first blush, it would be easy to say that searches from the lab and the field are not compatible given the vast differences between the manner in which search research is typically conducted in the lab and how search is performed in the field. However, such a conclusion would be both pragmatically unfortunate and empirically premature. Four significant hurdles were discussed in this chapter, and while each raises a critical concern when attempting to use lab-based findings to improve searches in the field, all show that with proper consideration, they can be overcome.

Target prevalence. The overwhelming majority of published cognitive psychology studies on visual search have employed paradigms in which targets appear on a substantial percentage of the trials. Yet, many field-based searches rarely have a target present (e.g., there is not a gun in the majority of baggage X-rays). While this difference initially presented itself as a critical hurdle for translating between the lab and the field, a number of studies have now explicitly focused on the impact of target prevalence (both in

cognitive psychology searches and in radiological searches). The hope is that, with careful experimentation, the effects of target prevalence will be isolated so that researchers can explore this topic for its own sake, but can also explore field-related visual search questions without prevalence serving as a confound. Several recent studies have made significant advances along this front, and target prevalence may no longer serve as a critical hurdle.

Number of targets and target categories. As for target prevalence, most cognitive psychology studies have employed search arrays that only have one possible target at any given time. Likewise, most lab-based search experiments inform the searcher of their exact target prior to the start of the search. In contrast, most searches in the field can have an unconstrained number of targets and targets from multiple different categories. This difference is potentially devastating given that the possibility of multiple targets within a search array could have broad influences on search strategies and accuracy. However, recent efforts have explicitly married lab- and field-based searches (e.g., Fleck et al., 2010), finding numerous commonalities with multiple-target visual searches. This not only suggests that field-based visual search can learn from searches in the lab, but, that even more so, career searchers can partner with cognitive psychologists to take advantage of the benefits afforded by testing in the lab. It is not easy to experiment with working radiologists or X-ray operators, so anything that can be tested out in the lab, such that it will translate, can be profoundly helpful. With the knowledge that multiple-target search in the lab can translate to the field, researchers are now equipped to use these search paradigms to explore the additional differences between the lab and the field.

Motivation and anxiety. The contexts in which searches are performed in the lab and the field are extraordinarily different, given both added motivation and added anxiety when conducting life-critical searches. In examining the effects of contextual motivation and anxiety in the lab, it becomes especially apparent why the deviations from standard search paradigms must be employed in order to properly investigate the effects of these factors in the field. In a multiple-target search paradigm—in which some trials only have one target, but other trials have more than one target— there were no differences in performance on single-target trials in motivated vs. non-motivated conditions (Clark et al., 2011) and in anxious vs. non-anxious conditions (Cain et al., 2011). In both cases, the differences were only apparent in dual-target trials. As is known from the investigation of multiple-target search in general, complex mechanisms may be interacting to cause performance differences that basic single-target searches are simply not sensitive enough to show. Searches in the field can contain more than one target and can be conducted in motivated and/or anxious contexts; by examining the effects of these factors in tandem, meaningful conclusions can be made about how these contexts may affect performance on field searches.

Levels of experience. Finally, experts have demonstrated vastly different performance on a variety of tasks when compared with inexperienced searchers. Even when the results may appear similar between the two groups, more sensitive analyses reveal that expert searchers could be approaching the tasks very differently. In these cases, one must be careful not to jump to conclusions about expert searchers from what is observed with undergraduate, inexperienced searchers. Other tasks, however, may

demonstrate similar processes between the two groups. Early work used real radiographs and actual radiologists to explore performance in the lab, but since inexperienced searchers do not have the experience to identify targets in such stimuli, researchers must use simplified displays accessible to both groups. By using these tasks, researchers can test both inexperienced and expert searches, and if in some tasks, performance is similar between the groups, conduct follow-up experiments using inexperienced searchers and potentially draw conclusions about experts from these data.

Despite the tremendous differences between lab and field searches, these hurdles are not insurmountable. Experimental search research is invaluable to the applied world, but only with an acknowledgement of the differences and shortcomings. By modifying the parameters of search tasks in the lab to account for the differences in target distributions in the field, by adding contextual factors present in the field such as motivation and anxiety to tasks in the lab, and by exploring the differences in performance between inexperienced and expert searchers, researchers are able to appropriately examine visual search processes as they exist in the applied world. While these are lofty requirements, ideally all examined simultaneously, with careful experimentation, we can understand the contributions of the individual factors and how they may interact. By accounting for all of these differences, we have the ability to use the performance of Tyler, the undergraduate, to inform and improve work conditions for Olivia, the X-ray operator.

References

Alvarez, G. A., & Cavanagh, P. (2004) The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*, 106-111.

Anderson, B.A., Laurent, P.A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*, 10367-10371.

Appelbaum, L. G., Schroeder, J. E., Cain, M. S., & Mitroff, S. R. (2011). Improved visual cognition through stroboscopic training. *Frontiers in Psychology, 2:276*. doi: 10.3389/fpsyg.2011.00276

Appelbaum, L. G., Cain, M. S., Schroeder, J. E., Darling, E. F., & Mitroff, S. R. (under revision). Stroboscopic visual training improves information encoding in short-term memory.

Ball, K., & Sekuler, R. (1982). A Specific and enduring improvement in visual motion discrimination. *Science, 218,* 697-698.

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research, 27,* 953-967.

Beard, B. L., Levi, D. M., & Reich, L. N. (1995). Perceptual learning in parafoveal vision. *Vision Research, 35(*12*)*, 1679–1690.

Bende, M. & Nordin, S. (1997) Perceptual learning in olfaction: professional wine tasters versus controls. *Physiology & Behavior, 62*, 1065–1070.

Bennett, S., Ashford, D., Rioja, N., & Elliott, D. (2004). Intermittent vision and one-hand catching: The effect of general and specific task experience. *Journal of Motor Behavior, 36,* 442-449.

Berbaum, K. S., Franken Jr., E. A., Dorfman, D. D., Miller, E. M., Krupinski, E. A., & Kreinbring, K. (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., Caldwell, R. T., & Lu, C. H. (2005). Can order of report prevent satisfaction of search in abdominal contrast studies? *Academic Radiology, 12,* 74-84.

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, E.A. Jr., Dorfman, D.D., Miller, E.M., Caldwell, R.T., Kuehn, D.M., & Berbaum, M.L. (1998). Role of faulty visual search in the satisfaction of search effect in chest radiography. *Academic Radiology*, *5*, 9-19.

Berbaum, K. S., Brandser, E. A., Franklin, E. A. J., Dorfman, D. D., Caldwell, R. T., & Krupinski, E. A. (2001). Gaze dwell times on acute trauma injuries missed because of satisfaction of search. *Academic radiology, 8*(4), 304-314.

Berbaum, K.S., Franklin, E.A., Caldwell, R.T., & Schartz, K.M. (2010). Satisfaction of search in traditional radiographic imaging. *The handbook of medical image perception and techniques*, 107-138.

Berlin, L. (1994). Reporting the ''missed'' radiologic diagnosis: Medicolegal and ethical considerations. Radiology, *192*, 183–187.

Biederman, I. & Shiffrar, M. (1987). Sexing day-old chicks: a case study and expert systems analysis of a difficult perceptual learning task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 640-645.

Bolfing, A., Halbherr, T., & Schwaninger, A. (2008). Using speed measures to predict performance in x-ray luggage screening tasks. *Proceedings of the IEEE International Carnahan Conference on Security Technology*, Zurich, October 5-8, 2009

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129(3),* 387–398.

Buck, L. Reaction time as a measure of perceptual vigilance. *Psychological Bulletin,*1966, 65, No. 5, 291-304.

Bond, A. B. (1983). Visual search and selection of natural stimuli in the pigeon: The attention threshold hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 292-306.

Cain, M. S., Adamo, S. H., & Mitroff, S. R. (2012, May). What eye-tracking can tell us about multiple-target visual search. Poster presented at the annual meeting of the Vision Sciences Society, Naples,FL.

Cain, M. S., Vul, E., Clark, K., & Mitroff, S. R. (in press). A Bayesian optimal foraging model of human visual search. *Psychological Science*.

Cain, M. S., Dunsmoor, J. E., LaBar, K. S., & Mitroff, S. R. (2011). Anticipatory anxiety

hinders detection of a second target in dual-target search. *Psychological Science, 22,* 866-871. doi:10.1177/0956797611412393

Cain, M. S., Landau, A. N., & Shimamura, A. P. (2012). Action video game experience reduces the cost of switching tasks. *Attention, Perception, & Psychophysics*. *74*(4), 641–647.

Cain, M. S., Vul, E., Clark, K., & Mitroff, S. R. (in press). A Bayesian optimal foraging model of human visual search. *Psychological Science*.

Callan, D. E., & Schweighofer, N. (2008). Positive and negative modulation of word learning by reward anticipation. *Human Brain Mapping, 29*: 237-249.

Camerer, C. F., & Hogarth, R. (1999). The Effects of Financial Incentives in Economics Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty*, 7-42.

Caplovitz, G. P., & Kastner, S. (2009). Carrot sticks or joysticks: Video games improve vision. *Nature Neuroscience, 12*(5)*,* 527–528.

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.

Chan, A. H. S., & Chan, C. Y. (2000). Validating the random search model for a double-target search task. *Theoretical Issues in Ergonomics Science*, *1*(2), 157-167. doi:10.1080/14639220050171315

Chan, A. H. S., Courtney, A. J., & Ma, C. W. (2002). Visual performance on detection tasks with double-targets of the same and different difficulty. *Ergonomics*, *45*(13), 934-948. doi:10.1080/00140130210166087

Chan, H. S., & Courtney, A. J. (1995). Visual performance on detection tasks with two targets. *International Journal of Human Factors in Manufacturing*, *5*(4), 417–428.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology 4, 55-81.*

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.

Chisholm, J. D., Hickey, C., Theeuwes, J., & Kingston, A. (2010). Reduced attentional capture in action video game players. *Attention, Perception, & Psychophysics, 72(3),* 667–671.

Clark, K., Samei, E., Baker, J., & Mitroff, S. R. (2011). Expertise in radiological screening and satisfaction of search. Poster presented at the annual Object Perception, Attention, and Memory meeting. Seattle, WA.

Clark, K., Cain, M. S. Adcock, R. A., & Mitroff, S. R. (2011). Interactions between reward, feedback, and timing structures on dual-target search performance. Poster presented at the annual meeting of the Vision Sciences Society, Naples, FL.

Clark, K., Fleck, M. S., & Mitroff, S. R. (2011). Enhanced change detection performance reveals improved strategy use in avid action video game players. *Acta Psychologica, 136*, 67-72.

Colzato, L. S., van Leeuwen, P. J. A., van den Wildenberg, W. P. M., & Hommel, B. (2010). DOOM'd to switch: Superior cognitive flexibility in players of first person shooter games. *Frontiers in Cognition*, *1*. doi:10.3389/fpsyg.2010.00008

Davis, M., & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry, 6*, 13-34.

Donohue, S. E., Woldorff, M. G., & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics, 72,* 1120–1129.

Dorval, M., & Pepin, M. (1986). Effect of playing a video game on a measure of spatial visualization. *Perceptual and Motor Skills, 62*, 159–162.

Dye, M. W. G., Green, C. S., & Bavelier, D. (2009). Increasing Speed of Processing With Action Video Games. *Current Directions in Psychological Science*, *18*(6), 321-326.

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.

Davis, M., & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry, 6*, 13-34.

De Lisi, R., & Cammarano, D. M. (1996). Computer experience and gender differences in undergraduate mental rotation performance. *Computers in Human Behavior, 12*(3), 351–361.

De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *The Journal of Generic Psychology, 16*(3), 272–282.

Eckstein, M. (2011). Visual search: A retrospective. *Journal of Vision, 11*(5): 14, 1-36.

Egglin, T. K., & Feinstein, A. R. (1996). Context bias: a problem in diagnostic radiology. *The Journal of the American Medical Association, 276*, 1752-1755.

Eysenck, M. W., Derakshan, N., Santos, R. & Calvo, M. G. (2007) Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*(2), 336-353.

Fahle, M. & Edelman, S. (1993). Long term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research, 33,* 397-412.

Fanselow, M. S. (1980). Conditional and unconditional components of post-shock freezing. *Pavlov J Biol Sci, 15*, 177.

Fiorentini, A., & Berardi, N. (1981). Learning in grating waveform discrimination: Specificity for orientation and spatial frequency. *Vision Research, 21(7),* 1149-1158.

Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science, 18*, 943-947.

Fleck, M.S., Samei, E., & Mitroff, S.R. (2010). Generalized 'satisfaction of search': Adverse influences on dual-target search accuracy. *Journal of Experimental Psychology: Applied, 16,* 60-71.

Franconeri, S. L., Hollingworth, A., & Simons, D. J. (2005). Do new objects capture attention? *Psychol Sci, 16*(4), 275-281.

Franken, E.A. Jr., Berbaum, K.S., Lu, C.H., Kanam, Dorfman, D.D., Warnock, N.G., Simonson, T.M., & Pelsang, R.E. (1994). Satisfaction of search in the detection of

plain-film abnormalities in abdominal contrast studies. *Investigative Radiology, 4,* 403-9.

Godwin, H. J., Menneer, T., Cave, K. R., & Donnelly, N. (2010). Dual-target search for high and low prevalence X-ray threat targets. *Visual Cognition, 18*(10), 1439-1463.

Gould, J. D., & Carn, R. (1973). Visual search, complex backgrounds, mental counters, and eye movements. *Perception and Psychophysics*, *14*, 125-132.

Gur, D., Rockette, H. E., Armfield, D. R., Blachar, A., Bogan, J. K., & Brancatelli, G. (2003). The prevalence effect in a laboratory environment. *Radiology, 228*, 10-14.

Gür, E., Turhan, P., Can, G., Akkus, S., Sever, L., Güzelöz, S., Çifçili, S. & Arvas, A. (2004). Enuresis: Prevalence, risk factors and urinary pathology among school children in Istanbul, Turkey. *Pediatrics International*, 46, 58–63.

Gur, D., Bandos, A. I., Fuhrman, C. R., Klym, A. H., King, J. L., & Rockette, H. E. (2007). The prevalence effect in a laboratory environment: changing the confidence ratings. *Acad Radiol*, *14*(1), 49–53.

Green, C. S., & Bavelier, D. (2003). Action video game play modifies visual selective attention. *Nature, 423*, 534–537.

Green, C. S., & Bavelier, D. (2006a). Effects 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(1),* 217–245.

Green, C. S., & Bavelier, D. (2007). Action video game experience alters the spatial resolution of vision. *Psychological Science, 18*, 88–94.

Griffith, J. L., Voloschin, P., & Gibb, G. D. (1983). Differences in eye-hand motor coordination of video-game users and non-users. *Perceptual and Motor Skills, 57(1),* 155–158.

Grillon, C., Baas, J. P., Lissek, S., Smith, K., & Milstein, J. Anxious responses to predictable and unpredictable aversive events. Behavioral Neuroscience. 2004;118:916–924.

Grillon, C., Morgan, C. A., Ill, Davis, M., & Southwick, S. M. (1998). Effects of experimental context and explicit threat cues on acoustic startle in Vietnam veterans with posttraumatic stress disorder. *Biological Psychiatry, 44,* 1027-1036.

Hickey, C., & Theeuwes, J. (2008). ERP correlates of inter-trial effects in visual search. *Journal of Vision*, 8, 6 (Abstract).

Hubert-Wallander, B., Green, C. S., & Bavelier, D. (2010). Stretching the limits of visual attention: The case of action video games. *WIREs Cognitive Science, Wiley, 1*, 1–9.

Kaplan, I. T., & Carvellas, T. (1965). Scanning for multiple targets, *Perceptual and Motor Skills, 21*, 239-243.

Karle, J. W., Watter, S., & Shedden, J. M. (2010). Task switching in video game players: Benefits of selective attention but not resistance to proactive interference. *Acta Psychologica, 134*(1), 70–78.

Karni, A. & Sagi, D. (1991) Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences USA, 88*, 4966-4970.

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature, 365,* 250-252.

Kiss, M., Driver, J., & Eimer, M. (2009). Reward priority of visual target singletons modulates event-related potential signatures of attentional selection. *Psychological Science, 20*, 245–251.

Klein, R. M., Gadbois, S. & Christie, J. J. (2004). Perception and detection of counterfeit currency in Canada: Note quality, trainng and security features. *Proc. SPIE,* V. 5310, Rudolf L. van Renesse; (Ed), Optical Security and Counterfeit Deterrence Techniques V. 1-12.

Kokich, V. O., Jr., Kiyak, H. A., & Shapiro, P. A. (1999). Comparing the perception of dentists and lay people to altered dental esthetics. *Journal of Esthetic Dentistry*, *11*, 311-324.

Koopman, B. O. (1956a). The theory of search: I. Kinematic bases. *Operations Research, 4*(3), 324-346.

Koopman, B. O. (1956b). The theory of search: II. Target detection. *Operations Research, 4*(5), 503-531.

Koopman, B. O. (1957). The theory of search: III. The optimum distribution of search effort. *Operations Research*, *5*(5), 613-626.

Körner, C., & Gilchrist, I. D. (2008). Memory processes in multiple-target visual search. *Psychological Research*, *72*, 99–105. doi:10.1007/s00426-006-0075-1

Kristjánsson, Á., Sigurjónsdóttir, Ó., & Driver, J. (2010). Fortune and reversals of fortune in visual search: Reward contingencies for pop-out targets affect search efficiency and

target repetition effects. *Attention, Perception & Psychophysics, 72*, 1229–1236.

Krueger, L. E., & Shapiro, R. G. (1980). Repeating the target neither speeds nor slows its detection: Evidence for independent channels in letter processing. *Perception & Psychophysics,* 28, 68-76.

Kundel, H. L. (1989). Perception errors in chest radiology. *Sem Resp Therapy*, *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.

Lang, P. J., Bradley, M. M. (2009). Emotion and the motivational brain. *Biological Psychology, 84*(3), 437-450.

Lau, J. S., Huang, L. (2010). The prevalence effect is determined by past experience, not future prospects. *Vision Research, 50*(15), 1469–1474.

Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience, 12,* 549–551.

Li, F-H., Cao, B-H., Xiao, F., & Li, H. (2011). The role of inhibitory control in the process of rare target detection. *Acta Psychologica, 43*(5), 509-518.

Li, H., Chan, J. S. , Cheung, S. Y., & Yan, J. H. (2012). Inhibitory control differentiates *rare target search* performance in children. *Percept Mot Skills*, *114*(1), 339-351.

Libera, C. D., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science, 17*(3), 222-227.

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.

Madden, D. J. (2007). Aging and Visual Attention. *Current Directions in Psychological Science. 16*(2)*,* 70-74.

Mackworth, N. H. (1950). Researches on the measurement of human performance. *Medical Research Council Special Report, London, Series 268.*

Matthews, N., & Welch, L. (1997). Velocity-dependent improvements in single-dot direction discrimination. *Perception & Psychophysics, 59*(1), 60–72.

McCarley, J. S., & Steelman, K. S. (2006). Elements of human performance in baggage x-ray screening. Proceedings of the 4th International Aviation Security Technology Symposium, Washington, DC.

Metlay, W., Sokoloff, M., & Kaplan, I. T. (1970). Visual search for multiple targets. *Journal of Experimental Psychology, 85*, 148-150.

Mitroff, S. R., & Hariri, A. (2010). Identifying Predictive Markers of Field Performance: The Potential Role of Individual Differences in Threat Sensitivity. *Institute for Homeland Security Solutions Research Brief.* https://www.ihssnc.org

Mitroff, S. R., Friesen, P., Bennett, D., Yoo, H., & Reichow, A. (revised manuscript under review). Enhancing ice hockey skills through stroboscopic visual training.

Mitroff, S. R., Biggs, A. T., Cain, M. S., Darling, E. F., Clark, K., Adamo, S. H., & Dowd, E. W. (2012). Visual search at the airport: Testing TSA officers. Poster presented at the annual meeting of the Vision Sciences Society, Naples, FL.

Menneer, T., Barrett, D.J.K., Phillips, L., Donnelly, N., & Cave, K.R. (2004). Search efficiency for multiple targets. *Cognitive Technology, 9*, 22-25.

Menneer, T., Stroud, M.J., Cave, K.R., Donnelly, N., & Rayner, K. . (2008). Eye movements in search for multiple targets. In Cognitive and Cultural Influences on Eye Movements, K. Rayner, D. Shen, X. Bai, and G. Yan (eds), Tianjin People's Press/Psychology Press.

Menneer, T., Cave, K. R., & Donnelly, N. (2009). The cost of search for multiple targets: effects of practice and target similarity. *Journal of Experimental Psychology: Applied, 15*(2), 125-139.

Menneer, T., Barrett, D. J. K., Phillips, L., Donnelly, N., & Cave, K. R. (2007). Costs in searching for two targets: dividing search across target types could improve airport security screening. *Applied Cognitive Psychology, 21*(7), 915-932.

Navalakkam, V., Koch, C., & Perona, P. (2009). Homo Economicus in Visual Search. *Journal of Vision*, *9*(1), 31, 1-16.

Murayama, K., & Kuhbandner, C. (2011). Money enhances memory consolidation – but only for boring material. *Cognition, 119*, 120-124.

Murty, V. P., LaBar, K. S., Hamilton, D. A., Adcock, R. A. (2011). Is all motivation good for learning: Dissociable influences of approach and avoidance motivation in declarative memory. *Learning and Memory*, *18*, 712-717.

Nakayama, K., & Martini, P. (2010). Situating visual search. *Vision Research*. doi: 10.1016/j.visres.2010.09.003

Neider, M. B., Boot, W. R., & Kramer, A. F. (2010). Visual search for real world targets under conditions of high target-background similarity: Exploring training and transfer of training in older adults. *Acta Psychologica*, *134*, 29-39.

Neisser, U. (1963). Decision time without reaction time: Experiments in visual scanning. *American Journal of Psychology*. New York: Appleton-Century-Crofts.

Neisser, U., Novick, R., & Lazar, R. (1963). Searching for ten targets simultaneously. *Perceptual and Motor Skills, 17*, 955-961.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Nelson, R., & Strachan, I. (2009). Action and puzzle video games prime different speed/accuracy tradeoffs. *Perception, 38*(11), 1678–1687.

Nodine, C. F., & Kundel, H. L. (1987). The cognitive side of visual search in radiology. In J.K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 573-582). North Holland: Elsevier Science Publishers B. V.

Ohl, R., Roedel, A., Binder, E., & Holsboer, F. (2003). Impact of high and low anxiety on cognitive performance in a modified hole board test in C57BL/6 & DBA/2 mice. *European Journal Neuroscience, 17*(1), 128-136.

Okagaki, L., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of Applied Developmental Psychology, 15*(1), 33–58.

Orosy-Filders, C., & Allan, R. W. (1989). Psychology of computer use: XII. Video-game play — Human reaction-time to visual-stimuli. *Perceptual and Motor Skills, 69*(1), 243–247.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40,* 1227-1268.

Parasuraman, R., & Davies, D. R. (1976). Decision-theory analysis of response latencies in vigilance. *Journal of Experimental Psychology: Human Perception and Performance, 2*(4), 578-590.

Poggio, T., Fahle, M., & Edelman, S. (1992). Fast Perceptual Learning in Visual Hyperacuity. *Science, 256*(5059), 1018-1021.

Poulton, E. (1890). The colours of animals: Their meaning of object recognition and use especially considered in the case of insects. London: Kegan Paul. 360 pp.

Quaiser-Pohl, C., Geiser, C., & Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and Individual Differences, 44*(3), 609–619.

Ramachandran, V. S., Braddick, O. (1973). Orientation-specific learning in stereopsis. *Perception, 2*(3), 371-376.

Renfrew, D.L., Franken, E.A. Jr., 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.

Ramachandran, V. S., Braddick, O. (1973). Orientation-specific learning in stereopsis. *Perception, 2*(3), 371-376.

Rhudy, J. L. & Meagher, M. W. (2000). Fear and anxiety: Divergent effects on human pain thresholds. *Pain, 84*(1), 65-75.

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.

Rosser, J. C., Lynch, P. J. Haskamp, L., Gentile, D. A., & Yalif, A. (2007). The impact of video games in surgical training. *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.

Saarinen, J., Levi, D. M. (1995). Perceptual learning in vernier acuity: What is learned? *Vision Research, 35*(4), 519–527.

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.

Schneider, W., & Shriffin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention, II: Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 1-66.

Schwaninger, A. (2003a). Reliable measurements of threat detection. *AIRPORT*, 01/2003, 22-23.

Schwaninger, A. (2003b). Evaluation and selection of airport security screeners. *AIRPORT*, 02/2003, 14-15.

Schwaninger, A. (2006a). Threat image projection: enhancing performance? *Aviation Security International, December 2006,* 36-41.

Schwaninger, A. (2006b). Liquid identification: Reacting to the terror threat. Analysing, controlling and adapting to meet new threats. *Airport, 5/2006,* 30-31.

Schwaninger, A. (2006c). Airport security human factors: From the weakest to the strongest link in airport security screening. *Proceedings of the 4th International*

Aviation Security Technology Symposium, Washington, D.C., USA, November 27 – December 1, 2006, 265-270.

Schwaninger, A. & Hofer, F. (2004). Evaluation of CBT for increasing threat detection performance in X-ray screening. In: K. Morgan and M. J. Spector, *The Internet Society 2004, Advances in Learning, Commerce and Security* (pp. 147-156). Wessex: WIT Press.

Schwaninger, A., & Wales, A.W.J. (2009). One year later: how screener performance improves in x-ray luggage search with computer-based training. *Proceedings of the Ergonomics Society Annual Conference 2009,* 381-389.

Schwaninger, A., Hardmeier, D., & Hofer, F. (2005). Aviation security screeners visual abilities & visual knowledge measurement. *IEEE Aerospace and Electronic Systems, 20*(6)*,* 29-35.

Sireteanu, R., & Rettenbach, R. (2000). Perceptual learning in visual searches generalizes over tasks, locations, and eyes. *Vision Research, 40*, 2902–2949.

Smith, T. Q., & Mitroff, S. R. (under review). Stroboscopic training enhances anticipatory timing.

Terlecki, M. S., & Newcombe, N. S. (2005). How important is the digital divide? The relation of computer and videogame usage to gender differences in mental rotation ability. *Sex Roles, 53*, 433-441.

Tinbergen, N. (1960). The natural control of insects in pine woods: Vol. I. Factors influencing the intensity of predation by songbirds. *Archives Neelandaises de Zoologie*, *13,* 265-343.

Townsend, J. T. (1990). Serial and parallel processing: Sometimes they look like Tweedledum and Tweedledee but that can (and should) be distinguished. *Psychological Science, 1,* 46-54.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97-136.

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. *Atten Percept Psychophys, 71*(3), 541-553.

Vogels, R., & Orban, G. A. (1985). The effect of practice on the oblique effect in line orientation judgments. *Vision Research, 25*(11), 1679–1687.

von Bastian, C. C., Schwaninger, A., & Michel, S. (2008). Do multiview X-ray systems improve X-ray image interpretation in airport security screening? *Zeitschrift für Arbeitswissenschaft*, *3*, 166-17.

Vreven, D., & Blough, P. M. (1998). Searching for one or many targets: Effects of extended experience on the runs advantage. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*, 98-105.

West, G. L., Stevens, S. S., Pun, C., & Pratt, J. (2008). Visuospatial experience modulates attentional capture: Evidence from action video game players. *Journal of Vision, 8*, 1-9.

Westheimer, G., McKee, S. P. (1978). Stereoscopic acuity for moving retinal images. *J.*

Opt. Soc. Am., 68, 450-455.

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.

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention*. (pp. 13-73). East Sussex, UK: Psychology Press.

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., Cave, K. R., Franzel, S. L. (1989). Guided Search: An Alternative to the Feature Integration Model for Visual Search. *Journal of Experimental Psychology: Human Perception and Performance,* 15(3), 419-433.

Wolfe, J. M. (1994). Guided Search 2.0: A Revised Model of Visual Search. *Psychonomic Bulletin & Review*, 1(2), 202-238.

Wolfe, J. M., & Gancarz, G. (1996). Guided Search 3.0: Basic and Clinical Applications of Vision Science, Dordrecht, Netherlands: Kluwer Academic, 189-192.

Wolfe, J. M., Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5,* 1-7.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature, 435*, 439-440.

Wolfe, J. M., Birnkrant, R. S., Kunar, M. A., & Horowitz, T. S. (2005). Visual search for transparency and opacity: Attentional guidance by cue combination? *Journal of*

Vision, 5, 257-274.

Wolfe, J. M., Van Wert, M. (2010). Varying Target Prevalence Reveals Two Dissociable Decision Criteria in Visual Search. *Current Biology*, *20*, 121-124.

Yantis, S., & Jonides, J. (1996). Attentional capture by abrupt onsets: New perceptual objects or visual masking. *J. Experimental Psychology: Human Perception and Performance, 22*(6), 1505-1513.

Yuji, H. (1996). Computer games and information-processing skills. *Perceptual and Motor Skills, 83*, 643-647.