Variation in Visual Search Abilities and Performance

by

Kait Clark

Department of Psychology & Neuroscience
Duke University

Date: ________________________

Approved:

___________________________
Stephen R. Mitroff, Supervisor

___________________________
Marty G. Woldorff

___________________________
R. Alison Adcock

___________________________
Elizabeth M. Brannon

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

2014
ABSTRACT

Variation in Visual Search Abilities and Performance

by

Kait Clark

Department of Psychology & Neuroscience
Duke University

Date: __________________________
Approved:

___________________________
Stephen R. Mitroff, Supervisor

___________________________
Marty G. Woldorff

___________________________
R. Alison Adcock

___________________________
Elizabeth M. Brannon

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

2014
Abstract

Visual search, the process of detecting relevant items within an environment, is a vital skill required for navigating one’s visual environment as well as for careers, such as radiology and airport security, that rely upon accurate searching. Research over the course of several decades has established that visual search requires the integration of low- and high-level cognitive processes, including sensory analysis, attentional allocation, target discrimination, and decision-making. Search abilities are malleable and vary in accordance with long-term experiences, direct practice, and contextual factors in the immediate environment; however, the mechanisms responsible for changes in search performance remain largely unclear. A series of studies examine variation in visual search abilities and performance and aim to identify the underlying mechanisms.

To assess differences associated with long-term experiences, visual search performance is compared between laypersons (typically undergraduates) and specific populations, including radiologists and avid action video game players. Behavioral markers of search processes are used to elucidate causes of enhanced search performance. To assess differences associated with direct practice, laypersons perform a visual search task over five consecutive days, and electrophysiological activity is recorded from the scalp on the first and last days of the protocol. Electrophysiological markers associated with specific stages of processing are analyzed to determine
neurocognitive changes contributing to improved performance. To assess differences associated with contextual factors, laypersons are randomly assigned to experimental conditions in which they complete a visual search task within a particular framework or in the presence or absence of motivation, feedback, and/or time pressure.

Results demonstrate that search abilities can improve through experience and direct training, but the mechanisms underlying effects in each case are different. Long-term experiences are associated with strategic attentional allocation, but direct training can improve low-level sensory analysis in addition to higher-level processes. Results also demonstrate nuanced effects of experience and context. On searches that contain multiple targets, task framework impacts accuracy for detecting additional targets after one target has been identified. The combination of motivation and feedback enhances accuracy for both single- and multiple-target searches. Implications for cognitive theory and applications to occupational protocols are discussed.
Contents

Abstract..............................................................................................................................................iv

List of Tables ........................................................................................................................................xiii

List of Figures .......................................................................................................................................xv

Acknowledgements .............................................................................................................................xvii

1. Introduction .....................................................................................................................................1

1.1. Overview of visual search ...........................................................................................................4

1.1.1. A brief history .......................................................................................................................5

1.1.2. Basics of search research .....................................................................................................6

1.1.3. Established influences on accuracy and response time .....................................................7

1.1.4. Theories of search ..................................................................................................................8

1.2. Visual search: Experience and learning ....................................................................................9

1.2.1. Radiologists: Professional searchers ....................................................................................10

1.2.1.1. Multiple-target search: A common source of errors ......................................................12

1.2.1.2. Comparing professional and non-professional populations ...........................................13

1.2.2. Action video game players and enhanced perception and attention ...............................15

1.2.2.1. Mechanistic explanations of superior cognition ...............................................................16

1.2.2.2. Prior and current investigations using visual search tasks ..............................................17

1.2.3. Laboratory-based practice in visual search .......................................................................18

1.2.3.1. Perceptual learning of basic sensory processes ...............................................................19

1.2.3.2. Neural mechanisms of visual search ................................................................................19
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.3. Neurocognitive mechanisms of improvement through visual search practice</td>
<td>20</td>
</tr>
<tr>
<td>1.3. Visual search: Context and motivation</td>
<td>22</td>
</tr>
<tr>
<td>1.3.1. Structure of search protocols and task framing</td>
<td>24</td>
</tr>
<tr>
<td>1.3.1.1. Search frameworks in radiology and airport security</td>
<td>24</td>
</tr>
<tr>
<td>1.3.1.2. Comparing the effects of task structure on multiple-target search performance</td>
<td>25</td>
</tr>
<tr>
<td>1.3.2. Impacts of motivation and related factors</td>
<td>27</td>
</tr>
<tr>
<td>1.3.2.1. Effects of motivation on visual search performance</td>
<td>28</td>
</tr>
<tr>
<td>1.3.2.2. Exploring interactions between motivation, feedback, and time pressure</td>
<td>29</td>
</tr>
<tr>
<td>2. Searching for multiple targets: How radiologists perform differently than non-professionals</td>
<td>31</td>
</tr>
<tr>
<td>2.1. Methods</td>
<td>35</td>
</tr>
<tr>
<td>2.1.1. Participants</td>
<td>35</td>
</tr>
<tr>
<td>2.1.2. Apparatus</td>
<td>36</td>
</tr>
<tr>
<td>2.1.3. Multiple-target search task</td>
<td>36</td>
</tr>
<tr>
<td>2.1.4. Temporal-order judgment task</td>
<td>38</td>
</tr>
<tr>
<td>2.2. Results</td>
<td>40</td>
</tr>
<tr>
<td>2.2.1. Data preparation and analyses</td>
<td>40</td>
</tr>
<tr>
<td>2.2.2. Accuracy</td>
<td>41</td>
</tr>
<tr>
<td>2.2.2.1. Overall accuracy</td>
<td>41</td>
</tr>
<tr>
<td>2.2.2.2. Single-target accuracy</td>
<td>42</td>
</tr>
</tbody>
</table>
2.2.2.3. Relative second-target accuracy ................................................................. 42
2.2.3. Response time ............................................................................................... 43
  2.2.3.1. Total trial time .......................................................................................... 43
  2.2.3.2. Individual target types ............................................................................. 44
  2.2.3.3. Target-detection order ............................................................................. 45
2.2.4. Control task (temporal-order judgment) results ........................................... 47
2.3. Discussion ......................................................................................................... 47
2.4. Acknowledgements ............................................................................................ 51

3. Enhanced change detection performance reveals improved strategy use in avid action
video game players .................................................................................................. 53
  3.1. Methods .......................................................................................................... 59
    3.1.1. Participants ................................................................................................. 59
    3.1.2. Apparatus and stimuli ................................................................................ 60
    3.1.3. Procedures .................................................................................................. 61
  3.2. Results ............................................................................................................. 64
  3.3. Discussion ........................................................................................................ 67
  3.4. Acknowledgements ............................................................................................ 70

4. Improvement in visual search with practice: Mapping learning-related changes in
neurocognitive stages of processing ........................................................................ 72
  4.1. Materials and Methods .................................................................................... 74
    4.1.1. Participants ................................................................................................. 74
    4.1.2. Search paradigm ......................................................................................... 75
    4.1.3. Behavioral data acquisition and analyses .................................................... 78
5.1.1. Participants ................................................................................................................. 103
5.1.2. Apparatus ...................................................................................................................... 103
5.1.3. Design ............................................................................................................................ 104
5.1.4. Scoring and feedback .................................................................................................... 106
5.1.5. Framing of task by condition ....................................................................................... 107
5.2. Results ............................................................................................................................... 109
5.2.1.Comparability of experimental parameters between conditions .......................... 109
5.2.2. Equivalent performance on basic measures ............................................................. 110
5.2.3. Dual-target accuracy and satisfaction of search ....................................................... 111
5.3. Discussion and conclusions ............................................................................................ 114
5.4. Acknowledgements .......................................................................................................... 119

6. Effects of motivation, feedback, and time pressure on multiple-target search performance ............................................................................................................................ 120
6.1. Methods ............................................................................................................................ 123
6.1.1. Participants .................................................................................................................... 123
6.1.2. Apparatus ...................................................................................................................... 124
6.1.3. Multiple-target search task ......................................................................................... 124
6.1.4. Conditions .................................................................................................................... 127
6.1.4.1. Motivation .................................................................................................................. 127
6.1.4.2. Feedback .................................................................................................................... 128
6.1.4.3. Time pressure ............................................................................................................ 128
6.2. Results ............................................................................................................................... 129
7.1.2.3. Expertise and enhanced cognition ................................................................. 153

7.1.3. Laboratory-based practice on a specific visual search task .............................. 155

7.1.3.1. Electrophysiological investigations of improvement in visual search ........ 156

7.1.3.2. Neural mechanisms underlying enhanced visual search ......................... 157

7.2. Context and motivation ................................................................................... 160

7.2.1. Task framing ............................................................................................... 162

7.2.1.1. Differences in SSMs between fixed objective and fixed duration conditions .............................................. 163

7.2.1.2. Interpretation of task framework effects .................................................. 164

7.2.1.3. Feasibility of implementing a fixed duration structure in occupational settings ......................................................................... 165

7.2.2. Dissociating effects of motivation, feedback, and time pressure .............. 166

7.2.2.1. SSM errors in all conditions .................................................................... 167

7.2.2.2. No effect of time pressure on multiple-target search accuracy ............. 168

7.2.2.3. The combination of motivation and feedback promotes multiple-target search accuracy ......................... 170

7.2.2.4. Implications for occupational protocols .................................................. 171

7.3. Closing thoughts ........................................................................................... 172

7.3.1. The nature of learning in visual search ...................................................... 173

7.3.2. Influences on multiple-target search errors ............................................... 174

References ............................................................................................................ 175

Biography ............................................................................................................. 194
List of Tables

Table 1: Means and standard deviations for single-target accuracy for radiologists and nonprofessionals................................................................. 41

Table 2: Means and standard deviations for single-target trial response times for radiologists and nonprofessionals................................................................. 43

Table 3: The time taken in seconds (with standard deviations) before identifying each target type for each trial type. For dual-target trials on which the high-salience target was found first, the response time for low-salience targets is the time between the identification of the high-salience target and the low-salience target. For all other measures, the reported response time is the time taken to identify the target type from the onset of the trial................................................................. 44

Table 4: Means of each measure, by group, with standard deviations in parentheses..... 65

Table 5: Means and standard deviations for (A) response time and (B) accuracy, before and after practice................................................................. 83

Table 6: (A) Means and standard deviations of peak amplitudes of the sensory-evoked N1 component, collapsed across left and right targets. (B) Means and standard deviations of latency of the peak of the N1 component in response to targets on the left and right sides of the display. ................................................................. 84

Table 7: (A) Means and standard deviations of the amplitudes of the N2pc component, collapsed across left and right targets (contralateral versus ipsilateral to the target popout). (B) Means and standard deviations of the latencies of the N2pc component, collapsed across left and right targets (contralateral versus ipsilateral to the target popout)..................................................................................... 86

Table 8: Means and standard deviations of the amplitudes of the SPCN component. .... 87

Table 9: Means and standard deviations of the onset latencies of the LRP component... 89

Table 10: Response Times by Trial Type. Means (and standard deviations) in seconds for each trial type in the Fixed Objective and Fixed Duration conditions. ........................................ 111

Table 11: Factors included in each experimental condition........................................ 127
Table 12: Means (and standard deviations) for accuracy for each relevant target type and trial type, SSM difference scores, false alarms, and response time (RT), by condition...131

Table 13: Means (and standard deviations) for all measures, collapsed across the time pressure factor. .......................................................... 134
List of Figures

Figure 1: Sample search array on a multiple-target search trial......................................................... 37

Figure 2: Percentage of dual-target trials wherein the high-salience or low-salience target was found first for radiologists and non-professionals. Data include dual-target trials on which both targets were detected. Error bars represent standard error of the mean. ...... 46

Figure 3: Sample cycle on a “change” trial. ........................................................................................................... 62

Figure 4: Hypothetical model demonstrating potential changes in (A) behavior and (B) and ERP components. (A) Response time is expected to decrease after practice. B) Horizontal arrows indicate potential latency shifts in the N1, N2pc, and LRP components. Vertical arrows indicate potential amplitude changes in the P1/N1, N2pc, and SPCN components. ........................................................................................................................................... 74

Figure 5: Sample stimulus display. Blue ellipses are distractors, the green ellipse is the relevant color popout target, and the red ellipse is the irrelevant color popout non-target. Participants respond as to the orientation of the green target ellipse. In this example, a participant would respond by pressing the button corresponding to “horizontal.” ................................................................................................................................... 76

Figure 6: Behavioral results for (A) response time and (B) accuracy, before and after practice. Response time decreased significantly over the course of practice, but there was no significant change in accuracy. .............................................................................................................. 84

Figure 7: (A) ERP traces of the sensory-evoked N1 component, collapsed across left and right targets, demonstrating an increase in amplitude after practice. (B) Distribution of N1-related activity over the scalp in response to targets on the left and right sides of the display.................................................................................................................................. 85

Figure 8: (A) ERP traces of activity used to calculate the differences waves for deriving the N2pc and SPCN components (contralateral versus ipsilateral to the target popout), collapsed across superior and inferior sites. (B) Difference waves displaying N2pc and SPCN components at the superior and inferior ROIs. (C) Distribution of N2pc-related activity over the scalp for Sessions 1 and 5. (D) Distribution of SPCN-related activity over the scalp for Sessions 1 and 5......................................................................................................................... 88

Figure 9: (A) ERP traces of activity used to calculate contralateral versus ipsilateral differences waves used for deriving the LRP component (contralateral versus ipsilateral
to the hand used for the motor response). (B) Difference waves displaying the LRP component. (C) Distribution of LRP-related activity over the scalp for Sessions 1 and 5.

Figure 10: Sample Trial. Example display for a dual-target trial. This display contains one high-salience target ‘T’ (far right, middle) and one low-salience target ‘T’ (middle, far bottom).

Figure 11: Sample Feedback Screen. Example feedback screen displayed following trial completion. On this example dual-target trial, the participant identified the high-salience target (denoted by the small circle) but missed the low-salience target (denoted by the large circle; red in experiment).

Figure 12: (A) Accuracy rates for low-salience targets in the Fixed Objective and Fixed Duration conditions: Single-target trials vs. dual-target trials (provided the high-salience target was detected first). Error bars represent standard error of the mean. (B) SOS difference scores (difference between accuracy rates in Figure (A) in the Fixed Objective and Fixed Duration conditions. Error bars represent standard error of the mean.

Figure 13: Accuracy Rates over Time. Accuracy rates for low-salience targets in single- and dual-target trials in the Fixed Objective and Fixed Duration conditions during each quarter of trials over the course of the experiment.

Figure 14: Sample array on a multiple-target search trial.
Acknowledgements

I would like to thank the many individuals who have contributed to this work – through mentorships, knowledge, ideas, support, and friendship.

The majority of this work is the product of a collaborative effort with my advisor, Steve Mitroff, who has been a tremendous source of inspiration and encouragement throughout my graduate career. Steve has exceeded every conceivable notion of the ideal advisor, serving as a role model, confidant, problem solver, and voice of reason. I thank Marty Woldoff, my secondary advisor, for his mentorship, his keen eye for detail, his refreshing (and often humorous) worldview, his rare nonjudgmental nature, and his persistence, and the other members of my committee, Alison Adcock and Liz Brannon. Alison has also played a crucial role in this work, offering her expertise related to motivation and cognition, and Liz has provided advice and suggestions essential to the theoretical organization of this work. An additional thanks goes to all of the members of the CCN for their support over the years as well as my undergraduate advisors, Pat Garrigan and John Jewell, for providing me with research opportunities early in my career.

Current and former members of the Mitroff lab have been incredible colleagues, including Adam Biggs, Matt Cain, Elise Darling, Emma Wu Dowd, and Betsy Schroeder, whose friendship and assistance have been invaluable, and a special thanks to Mat
Fleck. During my first year of graduate school, Mat generously extended his guidance and friendship, and much of this work was directly inspired by Mat’s graduate research. I am honored to follow Mat as Steve’s second graduate student. I also thank members of the Woldorff lab, especially Sarah Donohue and Joe Harris, and additional collaborators Greg Appelbaum, Tate Jackson, Ehsan Samei, and Jay Baker. A special thanks goes to the undergraduate students I have mentored, Clara Colombatto and Grace Kohut, for their enthusiasm and assistance.

I thank the members of my incoming graduate cohort, Drew Marticorena, Nick DeWind, Dave Barack, and Yoni Mazuz, and other members of the scientific community who have offered insight and friendship over the years, especially Nicole Petersen, Kristina Krasich, Deepu Murty, Sarah DeYoung, Nicole Codd, Evan Palmer, Kerry Jordan, Elissa Aminoff, Mina Cikara, Maital Neta, Felipe De Brigard, Sara Szczepanski, and Kara Blacker. And I thank countless friends outside of academia including Michelle Cantor, Natalie Trzcinski, Julie Klotz, Everett Codd, Brandi Gerew, David Beck, and Joe Gerew.

I thank my parents, Peggy and Dennis Clark, who made an enormous amount of sacrifices to provide me with the best educational opportunities available and encouraged me to ask questions and draw my own conclusions. I also thank Theresa and Joe (M.) Volosky for additional parental support over the recent years. Finally, I
offer a special thanks to Joe Volosky for his unwavering dedication, encouragement, perspectives, and insights.
1. Introduction

Cognitive processes form the crux of the human experience. In sensing and responding to a constantly changing environment, feeling a wide range of emotions, making complex decisions, and forming long-lasting memories, cognition comprises all we think and know. One vital cognitive task often taken for granted is visual search – an activity, broadly defined as the identification of target items among distractors, which encompasses a series of cognitive processes. Immersed in our ripe visual environments, we are constantly seeking to parse vast amounts of information into meaningful units, and it is easy to overlook the convenience of being able to accurately and efficiently identify any and all relevant items.

A recent review of visual search (Eckstein, 2011) opened with the claim, “Who searches? Everyone. Moreover, everyone searches all the time.” This seemingly bold statement is easily defensible when we pause to consider just how frequently we search. We search for our keys in our cluttered purses and for our cars in crowded parking lots. We search for our friends in a café and for a favorite latte on the menu. We search for files on our desktops, food in our refrigerators, and the list goes on. We conduct searches all the time, and most searches are relatively commonplace, but in some cases, visual searches can be critically important. For example, airport security screeners must identify harmful items in baggage, and radiologists must identify abnormalities in medical radiographs. Despite the ubiquitous nature of search and the fact that it is
sometimes life-or-death critical, human visual search is far from ideal – errors are often made, and searches are typically conducted for either too little or too much time. Thus, some fundamental questions emerge: How can we maximize search efficiency? What is the best way to increase both search speed and accuracy?

Basic research over several decades has investigated many factors that can impact visual search performance, establishing standards for extracting meaningful data and proposing explanations for cognitive limitations. Like many cognitive processes, visual search ability is malleable – search ability can change over time, and environmental factors can influence accuracy and efficiency. Visual search is cognitively complex, requiring the integration of multiple cognitive faculties, and understanding how these processes are modified is especially informative for cognitive theory. Additionally, accurate searching is critical for careers such as radiology and airport security, and insights gained from basic research can serve to improve selection, assessment, and training of personnel as well as work environments and procedures.

While a large body of research has examined factors associated with search performance, the underlying mechanisms are not always clear, and the implications of controlled lab-based studies may not be directly translatable to circumstances in career-based settings or to the searchers themselves. The goal of the research reported here is (1) to explore how visual search abilities and performance vary in accordance with long-term experiences and direct practice as well as environmental and contextual
factors, (2) to understand the mechanisms responsible for driving these changes, and (3) to discuss implications for both cognitive theory and its application to related career-based settings.

Chapter 1 provides a historical and theoretical overview of visual search and introduces the topics examined in the subsequent chapters. Chapters 2 through 6 each stand alone as empirical experiments but, collectively, contribute to a global understanding of the nature of cognitive flexibility. The first segment of studies (Chapters 2, 3, and 4) interrogates visual search performance in terms of changes that may occur over time, with respect to experience or learning. Chapter 2 focuses on performance differences between career searchers (radiologists) and laypersons, specifically with regard to multiple-target search. Chapter 3 compares avid action videogame players to non-videogame players on a change-detection task, which provides insight into potential mechanisms responsible for enhanced performance. Chapter 4 describes a training study in which electrophysiological data was recorded before and after training in order to determine the neural underpinnings of improvement on a visual search task. The second segment of studies (Chapters 5 and 6) examines visual search performance in terms of changes related to immediate, situational factors, such as context and motivation. Chapter 5 compares search performance among laypersons exposed to different instructional frameworks, akin to those employed in airport security and baggage screening. Chapter 6 examines the interactions between motivation, feedback, and time
pressure on visual search performance. Finally, Chapter 7 provides concluding thoughts on each study and on this body of work as a whole.

1.1. 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. The most basic visual searches operate via basic pattern matching; for example, detecting a green square within an array of red circles requires invoking only a green and/or square pattern template. More complex searches involve, for example, target items with variable appearances and/or orientations or the potential for multiple target items to be present within a single array. Successfully conducting a visual search relies upon a cascade of cognitive processes and may include perception (i.e., processing and interpreting visual features), attention (i.e., allocating resources to the relevant areas of a visual area), memory (i.e., storing a representation of the target item or items), and decision making (i.e., determining whether an ambiguous item is actually a target and deciding when to stop searching).

As such, search has been used extensively to learn about cognition. The study of visual search has made substantial contributions to 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), among others. Over the past several
decades, psychological research has made tremendous headway in understanding the processes contributing to the successful and efficient identification of target items as well as cognitive limitations related to search errors. However, many open questions remain about the flexibility of search abilities: How can search processes change over time? How and why do long-term changes occur? What environmental and/or contextual factors are associated with changes on a more immediate basis? Furthermore, how can the answers to these questions be applied to career-based searches in which optimal performance is especially critical?

1.1.1. A brief history

The first investigations of visual search were driven by its relevance to a primary evolutionary goal – 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. In the late nineteenth century, biologist Edward Poulton noted the variety of appearances within a single species of moth and speculated that this adaptive trait allowed for more effective evasion of predators (Poulton, 1890, p. 46). He proposed that it might be more difficult for predators to search for multiple types of wing patterns and colors, an idea later supported by research in cognitive psychology (e.g., Menneer, Barrett, Phillips, Donnelly, & Cave, 2007). Expanding on Poulton’s observations, Dutch zoologist Luuk Tinbergen found that insectivorous birds effectively maximized their rates of prey detection by restricting search to a few target types at a
given time – either the most common prey available or those that had been seen most recently (Tinbergen, 1960). In effect, his research demonstrated that non-human animals are sensitive to the statistics of their environments and able to quickly adapt strategies to maximize search efficiency, and contemporary work with human searchers has found similar results (e.g., Cain, Vul, Clark, & Mitroff, 2012). Pigeon studies have also illustrated that search is specialized for ecologically relevant tasks, as pigeons demonstrate a keen ability to find food (e.g., Bond, 1983) and effectively optimize their rates of food discovery.

While the work with non-human animals has clear parallels to human search processes, the first known studies specific to human visual search were conducted by a group of mathematicians during WWII and later published by Bernard Koopman (Koopman, 1956a; 1956b; 1957). Seeking to optimize procedures for detecting both enemy ships and lost personnel, the US Navy recruited Koopman and his team to identify factors associated with successes and failures in visual search. The ideas resulting from Koopman’s work, including the distribution of attention and criteria for termination, remain fundamentally important for current theories of search (e.g., Chun & Wolfe, 1996) and were empirically tested in the decades to follow.

1.1.2. Basics of search research

In a typical visual search experiment, participants are presented with an array of items with the goal of detecting or identifying a particular item, or “target,” among the
rest of the items, or “distractors.” In some experiments, participants are to identify the target (e.g., by making a localization mouseclick at the location of the target on the screen), and in others, participants simply respond whether the target is present or absent. Many studies have used response time as the main dependent variable by which performance is assessed, but accuracy has also been a focus.

Early studies of visual search sought to characterize human attentional processes and examined how observers could identify items as targets or non-targets without fully processing each item (Neisser, Novick, & Lazar, 1963) as well as mechanisms responsible for driving automatic and controlled processing (Schneider & Shiffrin, 1977). Thousands of subsequent studies from the 1980s to the present have explored how a variety of factors within the search display affect performance (see Nakayama & Martini, 2010; Eckstein, 2011, for reviews).

1.1.3. Established influences on accuracy and response time

Because of the necessary interaction between attentional and perceptual systems required for accurate searching, both the stimulus-driven salience of the targets (e.g., Einhäuser, Spain, & Perona, 2008; Foulsham & Underwood, 2008; Koch & Ullman, 1985; Masciocchi, Mihalas, Parkhurst, & Nieber, 2009) as well as top-down task demands and prior knowledge (e.g., Theeuwes, 2010) affect search efficiency. Investigations of factors contributing to basic perceptual stages have established how features contribute to bottom-up salience (Wolfe & Horowitz, 2004), and research on the attentional stage has
demonstrated how features capture attention (Yantis & Jonides, 1996; Franconeri et al., 2005), whether we attend to objects or locations (Goldsmith, 1998; Logan, 1996; Roelfema, Lammer, & Spekreijse, 1998; Yeari & Goldsmith, 2010), and how top-down control of selection is organized (Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004; Hamker, 2006; Theeuwes, 2010).

Other work has explored the effects of the nature of the display on performance, for instance, the eccentricity of the items (Wolfe & O’Neill, 1998), the characteristics of the background (Wolfe, Oliva, Butcher, & Arsenio, 2002), the visibility of the items to be searched (Wolfe, Birnkrant, Kunar, & Horowitz, 2005), additional cues in the environment (Fencsik, Urrea, Place, Wolfe, & Horowitz, 2006), the discriminability between targets items and distractors (Pashler, 1987). This large body of work has established the critical influences of many variables inherent in search displays and provides important standards to consider when designing the search paradigms used to explore broader influences on search processes.

1.1.4. Theories of search

In tandem with the large body of data that began to be amassed in the 1980s, two prominent theories emerged: feature-integration theory (Treisman & Gelade, 1980) and guided search (Wolfe, Cave, & Franzel, 1989; Wolfe, 2007). According to feature-integration theory, there are two distinct stages of visual search. First, the basic features of items (color, shape, orientation, etc.) are processed in the early stages of the visual
system, effortlessly and automatically. These features are organized into spatial maps such that, with directed attention, they can be bound together into integrated object percepts (Treisman & Gelade, 1980). A subset of these object percepts is then selected for further processing, which allows for more detailed analysis (e.g., determining whether an item is a target or a distractor).

The guided search model (e.g., Wolfe et al., 1989; Wolfe, 2007) has a similar, but less linear, conception of the processing stages involved in visual search. According to guided search, the basic features of items are used to direct the manner in which a viewer deploys attention across the display. Both basic sensory and selective attention processes are employed simultaneously, as basic perceptual processes identify the relevant features, and attention uses these features to guide the observer’s attention appropriately. While neither theory may completely account for the complexities of cognitive components required for visual search, both encompass a key element of integrated processing: Visual search relies upon the interaction between basic perceptual processes extracting simple features and directed attention conjoining these features to make sense of the environment.

1.2. Visual search: Experience and learning

Each cognitive component involved in visual search is malleable; abilities relating to perception, attention, memory, and decision making can change over time and be improved with training. Thresholds for basic perceptual discriminations can be
altered dramatically (e.g., McKee & Westheimer, 1978), visual attention and associated behavior can be improved (e.g., Beste, Wascher, Güntürkün, & Dinse, 2011), memory ability can be enhanced (e.g., Morrison & Chein, 2010), and decision-making processes can be trained (e.g., Catteeuw, Helsen, Gillis, & Wageman, 2010). It would follow, then, that visual search ability could be modified, and studies in which observers are directly trained on a visual search task have demonstrated such improvement and flexibility (e.g., Sireteanu & Rettenbach, 1995). While laboratory-based training studies offer the most direct means to explore how and why visual search processes (and resulting performance) may change, studying populations with long-term experiences associated with potential changes in visual abilities can also provide insight into how visual search performance varies.

1.2.1. Radiologists: Professional searchers

Visual search is a particularly relevant ability for several professional fields. Radiologists, among others (e.g., airport security officers, lifeguards), are tasked with conducting critical searches as the primary component of their jobs and spend years learning how to properly scan arrays to identify the appropriate target items. Generally, when compared to novices, trained professionals excel at tasks relevant to their fields of expertise; for example, expert chess players easily see patterns of moves (Chase & Simon, 1973), and wine connoisseurs are able to discriminate between many types of
wine (Bende & Nordin, 1997). Similarly, radiologists and cytologists are better able to
detect abnormalities in medical images than inexperienced searchers.

An important distinction is whether differences between experts and novices
emerge as a result of training or self-selection. When superior cognitive performance is
observed within an expert population, does it result from changes occurring during
training and practice, or are individuals who possess a certain ability simply more likely
to take up a hobby or enter a field? Some work has directly addressed this issue of
causality, finding that the experience itself leads to improved performance (e.g., farmers
improve their abilities to sort chickens by sex; Biederman & Shiffrar, 1987), but self-
selection may also play a role. Whether superior abilities among expert populations are
driven by experience, self-selection, or a combination of the two, such differences in
performance can offer insight into the cognitive mechanisms underlying successful
cognitive processes.

Radiologists comprise a unique expert population to study in this capacity, as
there is an existing body of work within the radiology literature that has examined
visual search processes and performance among practicing radiologists (see Berbaum,
2012, for a review). Because of the critical nature of radiological searches, wherein a
missed target could have fatal consequences, academic radiologists have sought to
pinpoint the common causes of search errors, and one particularly thorny source of
errors is the potential for multiple targets to be present within a single array.
1.2.1.1. Multiple-target search: A common source of errors

The problems associated with multiple-target searches became apparent in the radiological community when analyses of false negatives revealed that miss rates for abnormalities were substantially higher when there were more than one abnormality present within a radiograph. The idea that observers are more likely to miss a target after having identified one target in a display was named satisfaction of search (SOS; Smith, 1967), as it was originally believed to result from an early termination of search; that is, after having identified one target, an observer became “satisfied” with the meaning of the display and thus discontinued searching (Tuddenham, 1962). However, research has since ruled out early termination as a primary cause of SOS errors, as observers do continue to search after detecting one target (e.g., Berbaum, Franken, & Dorfman, 1991). Further investigations have identified alternative explanations for the marked decline in second-target accuracy, including attentional disruptions related to the identification of the first target and the depletion of available cognitive resources (Cain & Mitroff, 2013). Because the term “satisfaction of search” does not accurately reflect the cause of errors associated with multiple targets, researchers have recently shifted to use a broader descriptor, “subsequent search misses” (SSM; e.g., Adamo, Cain, & Mitroff, 2013)\. 

\(^1\)Throughout this document, “SOS errors” and “SSMs” are used interchangeably, as some chapters were published prior to the shift in nomenclature and employ the “SOS” terminology.
Not knowing how many targets might be present within a given search can dramatically alter how one goes about searching. Modern models of visual search that focus on single-target searches dedicate substantial efforts to predicting when a searcher should terminate a search (e.g., Wolfe et al., 2007), and these factors are exasperated in multiple-target search scenarios. Despite the fact that radiologists are explicitly informed of this pitfall and trained to attempt to avoid such errors, they still struggle with multiple-target searches, and related errors account for almost one-third of false-negative errors in some forms of radiological search (Anbari & West, 1997). As such, SOS errors, or SSMs, have been studied extensively within the radiological community (see Berbaum, 2012; for a review), and more recently, the phenomenon has been explored among non-professional populations. Non-professionals (e.g., undergraduates) also demonstrate a significant decline in accuracy for second targets when searching basic arrays (i.e., non-medical displays of simple shapes), suggesting that this error is not specific to radiologists and/or radiographs but reflects a more general cognitive limitation (e.g., Fleck, Samei, & Mitroff, 2010).

1.2.1.2. Comparing professional and non-professional populations

The errors associated with searching for multiple targets are observed across multiple populations and types of arrays, but are they arising for the same reasons? Related research within academic radiology has important implications for basic research in visual search and vice versa. However, direct translation of theories and
findings from one to the other can be troublesome because of the numerous differences between methodologies employed by basic and medical researchers. There are substantial discrepancies in the task parameters and stimuli (e.g., medical images vs. basic shapes) as well as in the demographics of the populations assessed (e.g., levels of motivation, age, education).

To begin to bridge the gap between these two productive areas of research, Chapter 2 presents a study that directly compares the performance of radiologists and non-professionals (undergraduates) on the same multiple-target search task. The particular task employed has reliably elicited SSMs among laypersons (e.g., Fleck et al., 2010), and radiologists are known to commit similar errors in practice. If SSMs result from a generalizable cognitive limitation, radiologists would likely exhibit such errors on a simplified search task as well. However, given the vast differences in search experience between radiologists and undergraduates, there are likely substantial differences group differences in search strategies and/or processes, and SSMs may arise for different reasons. The multiple-search task employed in Chapter 2 allows for the analyses of several behavioral markers over the course of searching, from which inferences may be drawn regarding the causes of SSMs. By comparing markers of processing between the populations and the implied accompanying strategies, this study aims to assess differences in search performance associated with radiological
experience and to determine the translatability between multiple-target search findings in cognitive laboratory and radiological settings.

1.2.2. Action video game players and enhanced perception and attention

While studying individuals with careers such as radiology can reveal the effects of experience (or, individual differences between groups), those who pursue specific hobbies and pastime activities can also exhibit unique cognitive abilities. One recreational pursuit that has been studied extensively within the past decade is action video game playing, specifically first-person shooter games. Individuals with extensive experience playing action videogames (videogame players; VGPs) tend to excel on a variety of perceptual and attentional tasks when compared to non-videogame players (NVGPs). Similar to career-based experiences, hobbies and pastime activities are uncontrolled and self-selected behaviors, but several studies have explored the causal role of videogame playing. Non-gamers trained on action video games have demonstrated similar improvements in performance as self-selected videogame players (e.g., 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 is an important mechanistic question of gamers’ benefits, but regardless of the causal nature of such benefits, differences between gamers and non-gamers have been consistently demonstrated.
1.2.2.1. Mechanistic explanations of superior cognition

VGPs (usually defined as those who play an action video games for more than six hours per week during the six months prior to testing) typically outperform NVGPs (who have played video games fewer than one hour per week during the six months prior to testing) on cognitive tasks related to attention and perception. For example, VGPs are able to respond more rapidly (Castel, Pratt, & Drummond, 2005; Dye, Green, & Bavelier, 2009) and switch between tasks faster (Cain, Landau, & Shimamura, 2012; Karle, Watter, & Shedden, 2010). VGPs also have improved spatial (Terlecki & Newcombe, 2005) and temporal (Donohue, Woldorff, & Mitroff, 2010) abilities and can enumerate briefly displayed items more quickly (Green & Bavelier, 2006b).

The cognitive mechanisms responsible for these differences, however, are unclear. Many tasks in which VGPs excel rely upon both basic perceptual skills and goal-oriented attention; do VGPs have heightened abilities in one or the other, or is their superior performance driven by enhancement in both? Some researchers have proposed a basic-sensory hypothesis, wherein videogame exposure hones perceptual abilities and allows for improved low-level vision (e.g., Dye et al., 2009; Green & Bavelier, 2006a; 2007; Li, Polat, Makous, & Bavelier, 2009; West, Stevens, Pun, & Pratt, 2008; Caplovitz & Kastner, 2009). Accordingly, VGPs may possess an increased capacity to process visual information relative to NVGPs. Alternatively, an improved-strategy hypothesis suggests that VGPs have enhanced higher-level abilities associated with attentional control and
shifts in attentional allocation (Cain et al., 2012; Chisholm, Hickey, Theeuwes, & Kingston, 2010; Hubert-Wallander, Green, & Bavelier, 2010), for generalized use across a variety of visually demanding tasks. In line with this account, VGPs 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).

1.2.2.2. Prior and current investigations using visual search tasks

Studies of visual search have often allowed for the dissociation of performance associated with basic sensory vs. attentional processes, and two studies have focused specifically on search abilities in VGPs (Castel et al., 2005; Hubert-Wallander, Green, Sugarman & Bavelier, 2011). Using a standard Posner cuing paradigm (e.g., Posner & Cohen, 1984), Castel et al. (2005) found no meaningful differences between VGPs and NVGPs in terms of search efficiency but found that VGPs simply respond more rapidly. When Hubert-Wallander and colleagues (2011) employed more difficult versions of the paradigm, with specific focuses on speed or accuracy, they reported superior search rates among VGP and suggested that differences potentially obscured by ceiling effects (e.g., Castel et al., 2005) could be revealed with more nuanced methodology. The authors did not find support for differences related to exogenous attention but proposed that the enhanced search efficiency of VGPs was driven by superior top-down attentional abilities.
However, a key component of many visual searches is self-directed attentional allocation within complex search arrays. Outside the constraints of a cuing task, do VGPs perform differently than NVGPs when freely viewing a scene with a specific goal, and if so, how? Chapter 3 explores performance over the course of searching during a change-detection flicker task. In this paradigm, observers are to attempt to identify a change between two rapidly presented scenes over several presentation cycles. In line with consistent cognitive benefits within VGPs, it is expected that VGPs will require fewer cycles to correctly identify the location of the change. Importantly, behavioral markers leading to eventual target detection allow for analyses of search processes and have implications for the dissociation between the basic-sensory and improved-strategy hypothesis discussed in Section 1.2.2.1.

1.2.3. Laboratory-based practice in visual search

Experiences such as extensive action video-game playing may enhance perceptual and attentional processing, resulting in generalizable benefits that facilitate visual search performance. In such cases, experience in one domain may lead to broad cognitive changes in related domains that rely on similar processes. These sorts of changes may be related to a global sort of “expertise” (see Chapter 7 for a more detailed discussion of expertise), but a related body of research has focused on identifying cognitive improvements associated with direct training or practice on a specific task.
1.2.3.1. Perceptual learning of basic sensory processes

Basic sensory abilities can change with practice, a phenomenon referred to as *perceptual learning*. 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). With extensive practice, participants improve in discrimination abilities for numerous basic features, including the orientation of a line (Ramachandran & Braddick, 1973; Fiorentini & Berardi, 1981; Matthews & Welch, 1997), direction of motion (Ball & Sekuler, 1982; 1987), and vernier acuity (Westheimer & McKee, 1978; Saarinen & Levi, 1995; Beard, Levi, & Reich, 1995). Such changes are often attributed to plasticity in the primary visual cortex (V1), and in the discrimination tasks described above as well as 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.

1.2.3.2. Neural mechanisms of visual search

In addition to basic perception, however, visual search requires higher-level cognitive processes, and performance on most visual search tasks would not improve with stimulus-specific sensory improvement alone. Activity in V1 may be responsible for a portion of the processing required for visual search, allowing for the creation of saliency maps in the early stages of processing (e.g., Li, 2002), but later stages of processing required for successful visual search rely on higher-level areas as well.
Research has identified several areas associated with visual search processing including V4 (Motter, 1994), the frontal eye fields (Schall & Hanes, 1993), the lateral intraparietal cortex (Toth & Assad, 2002), and the superior colliculus (e.g., McPeek & Keller, 2002), in which neurons selectively respond to target-specific features relative to features of distractors. Evidence from monkey physiology has demonstrated that neurons in low-level visual areas such as V1 respond to specific perceptual characteristics, in line with the perceptual learning work described above, but that neurons in areas of higher-level processing associated with visual search respond to target-specific features, regardless of their physical characteristics (e.g., Ptak, 2011). Plasticity in V1 is well established, but the nature of coding and representations in higher-level areas is less clear, as is the degree to which they can be trained.

1.2.3.3. Neurocognitive mechanisms of improvement through visual search practice

Research has established that search performance can be significantly enhanced through experience and practice (e.g., Sigman & Gilbert, 2000), and improvements are not specific to the trained stimulus (e.g., Sireteanu & Rettenbach, 1995), implying that improvements are occurring after the basic feature-analysis stage. However, given the numerous stages of processing following sensory analysis, which cognitive components are changed to allow for more efficient searching? Improved search performance could result from enhancements in various cognitive processing stages, including 1) sensory processing, 2) attentional allocation, 3) target discrimination, 4) motor-response
preparation, and 5) response execution. Human electrophysiological studies of visual search typically use time-locked event-related potentials (ERPs) to delineate the neural mechanisms of attentional processing, and research has identified specific ERP components that are reliably associated with the above stages involved in visual search processing, namely, 1) the *posterior visual N1*, a negative-polarity wave (latency ~150 ms) that reflects early sensory-evoked processing (Mangun & Hillyard, 1991); 2) the *N2pc* (negative-polarity posterior-contralateral, latency ~225 ms) associated with the shift of attention to a lateralized stimulus location (Luck & Hillyard, 1994a); 3) the *SPCN* (sustained posterior-contralateral negativity, latency 300-450 ms) or *CDA* (contralateral delay activity) that has been related to maintenance and manipulation of information in visual working memory (Jolicoeur et al., 2006; Vogel & Machizawa, 2004) and/or to cognitive processing required for target discrimination (Jolicoeur et al., 2008); and 4) the *motor-related LRP component* (lateralized readiness potential) that reflects the initiation of a motor response (Coles, 1988). The fifth stage, response execution, is marked by behavioral response time.

To determine which stages of processing are responsible for driving increased search efficiency, Chapter 5 presents research that compares these neurocognitive markers of processing before and after training on a visual search task. Participants completed a 5-day visual-search training protocol, and electrophysiological data was recorded on the first and fifth day. In line with prior studies (e.g., Sireteanu &
Rettenbach, 1995), behavioral response times are expected to decrease with training, and analysis of electrophysiological activity associated with each processing stage leading to the ultimate response allows for a comprehensive investigation of neurocognitive mechanisms underlying training in visual search.

1.3. Visual search: Context and motivation

While experience and learning are associated with differences in visual search performance over time, the environment in which an observer is searching can influence search performance more instantaneously. The impacts of situational factors, such as context and motivation, are especially critical when attempting to equate performance between different populations and draw conclusions about the causes of potential improvements. Career searchers (e.g., radiologists) carry the burden of awareness that an error in any search could result in fatalities, which may elicit an element of anxiety and alter search processes and performance. Presumably, professional searchers are also quite motivated to perform well, not only because of the life-saving potential of accurate searching, but also because of job security and/or pride in their lines of work. With the ultimate goal of translating findings from basic research for application in relevant occupational settings, it is crucial to understand the impacts of these factors on visual search performance.

Research has established reliable relationships between a variety of circumstantial factors (e.g., mood, fatigue) and performance across the spectrum of
cognitive tasks, from simple categorization to cognitive control (e.g., Nadler et al., 2010; van der Linden, Frese, & Meijman, 2014). Influences such as mood and fatigue are largely out of the control of employers and would be difficult to manipulate via workplace protocols, but other elements known to impact cognition (e.g., performance-based reward; Savine & Braver, 2010) can be implemented with the goal of enhancing performance.

Occupational searches tend to present complex challenges for searchers, including the possibility of multiple targets in a display (discussed in Section 1.2.1.1) and the fact that hazardous targets occur very infrequently. These complexities are not only inherent in radiological searches but also in searches conducted by airport security personnel. Furthermore, the unique challenges associated with multiple and/or rare targets tend to be especially sensitive to contextual manipulations. For instance, when searching for multiple targets under experimentally induced conditions of anxiety, observers suffer significant declines in accuracy on dual-target trials (i.e., SSMs) relative to a no-anxiety control condition (Cain, Dunsmoor, LaBar, & Mitroff, 2011). On single-target trials in the same experiment, however, there were no differences in performance between the conditions, suggesting that more cognitively taxing tasks may retain a heightened sensitivity to situational influences.
1.3.1. Structure of search protocols and task framing

The structure within which a series of searches are conducted defines, in part, contextual influences at play. If airport security personnel are required to spend lengthy periods of time at the checkpoint, does fatigue set in and result in accuracy declines? If there is a long line at the checkpoint, or hundreds of medical scans yet to be assessed, does this hefty workload elicit stress? If employees are rewarded for excellent performance or punished for subpar searching, how do reward and punishment impact cognition?

The possible structures of search series and the resultant cognitive impacts are innumerable, but the manner in which a task is framed can also have tremendous impacts on human cognition – even when all circumstantial factors are seemingly identical. For instance, a substantially larger percentage of respondents are likely to support a medical program if presented in terms of the proportion of lives saved rather than the proportion of lives lost (Tversky & Kahneman, 1981). In both cases, the percentage of survivals and deaths are identical, but the positive emotions elicited by the prospect of saving lives tends may result in a different decision-making process than when the same situation is presented in terms of the morbid side of the outcome.

1.3.1.1. Search frameworks in radiology and airport security

Both radiologists and airport security screeners routinely conduct series of complex searches, but they do so within different constraints: Radiologists typically
search with a fixed objective (e.g., assigned to assess 45 mammography images), while airport security screeners search for a fixed duration (e.g., scheduled to serve as an X-ray screener at the passenger checkpoint for a 30-min period). Both radiologists and airport security screeners are trained to maximize accuracy and, in effect, should be attempting the same process: carefully examining each display for potentially harmful targets, regardless of the number of cases yet to be scanned or the amount of time left before the end of a shift. Do these opposing frameworks have differential impacts on cognitive processing?

1.3.1.2. Comparing the effects of task structure on multiple-target search performance

To continue to elucidate the causes of multiple-target search errors, Chapter 6 presents a study in which two groups of laypersons complete a multiple-target search task within a structure similar to radiological searches (searching with a fixed objective) or a structure similar to airport security searches (searching for a fixed duration). The task was designed to be objectively identical for the two groups; participants are instructed to complete a specified number of accurate searches in as few minutes as possible (fixed objective) or to complete as many accurate searches as possible within a specified number of minutes (fixed duration). Speed and accuracy are equally critical in both cases, as one must balance searching accurately enough (to ensure trials are correct) and quickly enough (to ensure completion of enough trials).
A prior study demonstrated a relationship between time pressure and multiple-target search errors, wherein SSMs occurred more frequently when there was a trial time limit of 15 seconds than when the trial time limit was 30 seconds (Fleck et al., 2010). Importantly, participants rarely exceeded either time limit, and trial time was roughly equivalent between the two conditions (~10-11 seconds); as such, it is unlikely that the increase in SSMs observed in the 15-second condition resulted from an inability to complete the search (and therefore, increased likelihood to have not yet found a second target). Instead, the difference may have been driven by an impact of time pressure on participants, which could elicit anxiety – a factor shown to increase SSMs (Cain et al., 2011).

Given the relevance of multiple-target search errors to both radiology and airport security, as well as the impacts of time pressure and anxiety on second-target errors, the study discussed in Chapter 6 explores how the standard structures of search protocols within occupational settings differentially impacts search performance. Critically, neither of the search structures employed included a direct manipulation of anxiety or trial-based time pressure. There was no external mechanism of induced anxiety nor were there time limits for any individual trial. While there was likely an overall sense of “time pressure” in both groups, the experiment was self-paced, and the participants were free to find an optimal pace by which to maximize both speed and accuracy.
In the absence of trial-based time limits and experimentally induced anxiety, prior studies have not found substantial SSMs, and thus, if there are no contextual factors at play, neither group would be expected to commit SSMs. Alternatively, if SSMs occur within either group, or if differences emerge between the groups, the structure of the search tasks may be related to the induction of cognitive states in which observers are prone to error. Analyses will provide insight into the nature of task demands associated with search errors, further understanding of the impacts on multiple-target search, and implications for search structures within parallel professional settings.

1.3.2. Impacts of motivation and related factors

There is often great variability in visual search performance between participants, even when all participants are drawn from the same population (e.g., undergraduates). While individual differences in search ability likely contribute to the variability, some variability may also reflect 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.

When attempting to compare visual search performance between professional/experienced groups and laypersons, an important concern is whether there are differences in levels of motivation between the groups and the potential resulting
effects on performance. While controlled laboratory-based studies cannot achieve the type of motivation inherent in career-based searches (i.e., naïve participants cannot reasonably or ethically believe that lives are in their hands), one reasonable approximation for motivation is a performance-based monetary reward. The prospect of receiving money for good performance provides an effective global incentive that, for most people, will increase interest and effort (e.g., Camerer & Hogarth, 1999).

1.3.2.1. Effects of motivation on visual search performance

Several studies have employed a monetary reward to examine the impacts of motivation on visual search performance, finding impacts of motivation on attentional selection (e.g., Libera & Chelazzi, 2006; Kiss, Driver, & Eimer, 2009), priming (e.g., Hickey & Theeuwes, 2008; Kristjánsson, Sigurjónsdóttir, & Driver, 2010) and attentional capture (e.g., Anderson, Laurent, & Yantis, 2011). These studies present convincing evidence that monetary rewards can improve visual search performance; however, they focus primarily on changes in the speed of attentional deployment. While this is a critical component of search, career-based searches tend to place a larger emphasis on accuracy over speed.

As discussed, occupational searches are wrought with significant complexities – multiple targets exist with an array, and harmful targets (e.g., tumors in radiology, bombs in airport security) are exceedingly rare. Only a small percentage of radiographs contain a harmful abnormality, and only a small percentage of airport luggage items
contain malicious contraband. Target items tend to be missed more frequently when they occur rarely, an idea known as the prevalence effect (e.g., Wolfe, Horowitz, & Kenner, 2005; Wolfe et al., 2007; Fleck & Mitroff, 2007).

Because performance declines associated with the prevalence effect could be explained by declines in vigilance and loss of interest, a recent study investigated whether the prevalence effect could be overcome when participants were sufficiently motivated (Navalpakkam, Koch, & Perona, 2009). Participants searched for a target object in a cluttered scene, with levels of 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 with the prospect of reward (Navalpakkam et al., 2009).

1.3.2.2. Exploring interactions between motivation, feedback, and time pressure

How might errors associated with searching for multiple targets also be alleviated with motivation via a monetary reward? Further, how do motivational incentives interact with other factors associated with multiple-target search performance? Building upon the work discussed in Chapter 5, a series of experiments in
Chapter 6 allow for a comprehensive interrogation of the individual contributions of motivation, feedback, and time pressure on multiple-target search performance as well as the interactions between the factors. Time pressure is associated with declines in second-target accuracy (e.g., Fleck et al., 2010), but some results in Chapter 5 suggest that motivation and feedback could prevent SSM errors typically observed under conditions of time pressure. By examining multiple-target search performance in the presence or absence of each iterative combination of contextual factors, the experiments in Chapter 6 can delineate unique contributions of situational factors related to multiple-target search accuracy.
2. Searching for multiple targets: How radiologists perform differently than non-professionals

Professionals tend to exhibit superior abilities in their career domains relative to non-professionals. For example, professional musicians can better discriminate pitch than non-musicians (Kishon-Rabin, Amir, Vexler, & Zaltz, 2001), wine connoisseurs excel at discriminating between fine wines (Bende & Nordin, 1997), and bank tellers can better detect counterfeit currency (Klein, Gadbois, & Christie, 2004). For fields that demand high levels of accuracy, it is of particular interest to investigate how the refined skills of professional populations may differ from non-professionals. One domain in which accuracy is especially critical is that of visual search, the process of trying to find targets among distractors. Visual search is used in everyday life to identify relevant items and navigate the visual environment, and it is especially important for careers such as radiology and airport security screening, in which a missed target could have fatal consequences. Radiologists and airport security screeners undergo training regimens in preparation for conducting accurate searches of medical images and luggage, respectively, and perform these tasks repeatedly over the course of their careers.

Despite extensive training and hands-on experience, however, radiological and airport security searches are not without errors. Research has sought to identify how and why these searches may go awry, with some efforts focused specifically on the
visual abilities of radiologists (see Berbaum, 2012, for a review) and others on the search skills of airport security screeners (e.g., McCarley, Kramer, Wickens, Vidoni, & Boot, 2004; Gale, Mugglestone, Purdy, McClumpha, 2000). Collectively, these research efforts help to inform the broad goal of determining whether search failures can be alleviated with training and experience.

Comparing the search abilities of professionals and non-professionals can provide critical insight into how professionals perform well (and how they commit errors); however, such a comparison is not necessarily straightforward. Professional searchers may be more motivated to perform well than non-professionals or may spend a longer time searching, confounding improvements with motivational differences and a speed/accuracy tradeoff (see Biggs, Cain, Clark, Darling, & Mitroff, 2013). Additionally, non-professionals do not have the knowledge to identify specialized targets (e.g., tumors), providing professionals with an unfair advantage when comparing task-specific performance. To properly assess group differences, the experimental task must consist of target items that are identifiable by professionals and non-professionals alike, and differences related to motivation and speed must be considered.

Both professionals and non-professionals commit errors during visual search (e.g., Anbari & West, 1997; Fleck, Samei, & Mitroff, 2010), but relative differences between the groups are less clear. Understanding how the groups may search differently could inform the mechanisms underlying search as well as offer potential solutions for
improving performance. Relative to non-professionals, radiologists demonstrate increased perceptual sensitivity on visual discrimination tasks (e.g., Sowden, Davies, & Rolling, 2000); however, searching a complex scene for specific items requires additional cognitive faculties (e.g., attention, memory) beyond basic perception. A particular problem in radiology is the fact that medical images may contain more than one target (e.g., a fracture and a tumor), and the detection of a second target is less successful after the identification of one target in the same display. Originally termed “satisfaction of search” (SOS; Smith, 1967), the phenomenon was believed to arise from an early termination of search, as the observer became “satisfied” with his or her reading of the image after having identified one target (Tuddenham, 1962).

More recent investigations have ruled out early-termination as a primary cause of SOS errors, as observers do continue to search medical images after detecting one target (e.g., Berbaum, Franken, & Dorfman, 1991). Instead, the decline in second-target accuracy may result from faulty decision making (Berbaum, Franken, & Dorfman, 1998) or faulty pattern recognition (Samuel, Kundel, Nodine, & Toto, 1995). Accuracy for second targets may also suffer from the depletion of cognitive resources expended during the identification of one target, leaving subsequent cognitive processes more prone to error (e.g., Cain & Mitroff, 2012). Perhaps because of the added cognitive complexities associated with searching for multiple targets, second-target accuracy tends to be uniquely sensitive to contextual influences (e.g., anticipatory anxiety, Cain,
Dunsmoor, LaBar, & Mitroff, 2011; monetary incentives, Clark, Cain, Adcock, & Mitroff, 2011; and time pressure, Fleck et al., 2010) that affect single-target searches to a lesser extent or differently. Most research examining performance on multiple-target search tasks has employed real medical images as stimuli and career radiologists as participants (e.g., see Berbaum, 2012, for a review), but the SOS effect has also been found using simplified stimuli and non-professional searchers (e.g., Fleck et al., 2010; Cain et al., 2011; Cain & Mitroff, 2012). As such, SOS represents a generalizable cognitive phenomenon, occurring in both professionals and non-professionals, in both clinical and simplified experimental paradigms.

The goal of the current study is to compare the performance of professional and non-professional searchers on a simplified multiple-target search task to elucidate the nature of multiple-target search. Errors associated with searching for multiple targets occur in both radiology professionals and non-professionals, but it is unclear whether there are group differences in terms of search behaviors and whether the causes of errors can be generalized across populations. To date, only one study (Nodine & Krupinski, 1998) has compared professionals and non-professionals on a multiple-target search task. When searching non-radiological stimuli (line-drawn complex scenes), radiologists spent more time searching but were no more accurate than non-professionals. While the displays employed contained multiple targets, the analyses focused on general accuracy rather than performance related to searching for additional targets per se. Detailed
analyses of relative multiple-target search performance (e.g., Fleck et al., 2010) can reveal errors specific to multiple-target search that are not apparent with standard assessments of search accuracy. Given the specific sensitivity of second-target accuracy (e.g., Cain et al., 2011) and the emphasis on avoiding SOS errors within the radiological community, we examined differences in multiple-target search performance between radiologists and non-professionals.

2.1. Methods

2.1.1. Participants

Eight radiologists and 10 undergraduate students were recruited from the Duke University community. Participants in the Radiologist group were recruited through emails sent to the faculty and fellows of the Breast Imaging Division and to diagnostic radiology residents of the Duke University Medical Center Department of Radiology. They were asked to respond if they were willing to participate in a computer-based experiment for a visual cognition study. All were actively practicing radiologists or radiologists in training with 3 to 23 years of experience reading medical images (Mean age=36.5 years; SD=8.42; all female; Mean experience=9.5 years; SD=8.43; 2 residents, 3 fellows, 4 faculty). Because of the limited number of radiology participants available, there is not enough power to explore differences related to the more- and less-experience radiologists. Participants in the non-professional group were all undergraduates with no professional search experience (Mean age=22.2 years; SD=4.39; 8
female). All participants provided informed consent. Radiologists received $10 for participation, and non-professionals received either course credit or $10.

2.1.2. Apparatus

All participants completed the task in the Duke Visual Cognition lab with the same apparatus. Stimuli were presented on a Dell Inspiron computer with a 20-inch CRT monitor and programmed in MATLAB (The MathWorks, Natick, MA) using the Psychophysics Toolbox (Version 3.0.8, Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007). Participants were seated at a viewing distance of approximately 57 cm from the screen.

2.1.3. Multiple-target search task

Participants completed a multiple-target visual search task that has previously revealed an SOS effect (Fleck et al., 2010, Experiment 3). Each trial contained 25 items; the items consisted of target ‘T’ shapes and distractor pseudo-‘L’ shapes with a stroke width of 0.3°. The target ‘T’ shapes were comprised of a short bar (0.9° long) that approached a longer bar (1.3° long) at its exact midpoint, and the distractor pseudo-‘L’ shapes were comprised of a short bar that approached a longer bar at any location other than its exact midpoint. All items subtended a total area of 1.3° x 1.3° of visual angle and were presented on a rendered grayscale “cloud” background with a brightness range of 10%-50% black. Each trial contained 0, 1, or 2 targets, and participants were informed of the number of possible targets. Distractor pseudo-‘L’ shapes were always between 28%
and 66% black, and target ‘T’ shapes were presented in two levels of visibility: *high-saliency targets* (easy to spot; 66%-70% black) and *low-saliency targets* (hard to spot; 28%-40% black). Dual-target trials contained one high-saliency target and one low-saliency target. Each stimulus was situated within randomly selected cells of an invisible 8x7 grid, and the total stimulus space was 25.4° x 19.1° of visual angle (See Figure 1).

![Sample search array on a multiple-target search trial.](image)

Each trial began with a fixation cross appearing for 0.5 s at the center of the screen, which was replaced with a search array of 25 items on the cloudy background. Participants were instructed to use the mouse to click on each target ‘T’ they found and then to click a ‘DONE’ button at the bottom of the screen when they decided to
terminate their search. Participants were also provided with the option of correcting an accidental click or mistake by clicking a ‘CLEAR’ button at the bottom of the screen. After the completion of a trial, the cursor was reset in the center of the screen at the start of each new trial. There was a trial time limit of 15 seconds (based on an earlier study, Fleck et al., 2010) after which no further clicks were accepted, and if a participant exceeded this limit, a message appeared encouraging the participant to try to complete his or her search in the time allowed.

The experiment began with a brief practice session of 25 trials during which on-screen feedback was provided. The experimental session followed the practice session and consisted of 5 blocks of 51 trials each for a total of 255 trials. Four types of trials were employed throughout the experiment: no-target trials, single-target trials with a high-salience target, single-target trials with a low-salience target, and dual-target trials with one high-salience target and one low-salience target. Trial distribution was based upon Fleck et al. (2010) with rates of 20% no-target trials, 40% high-salience single-target trials, 16% low-salience single-target trials, and 16% dual-target trials.

2.1.4. Temporal-order judgment task

To address potential confounding group differences (e.g., level of motivation, age, education), all participants also completed a control task, wherein they made judgments about the order in which two squares appeared on a computer screen. This task was selected as a control task, as there was no a priori reason to assume that
participants with radiology experience would have a superior ability in judging the temporal appearance of items.

Stimuli were presented using the same equipment as in the SOS task but programmed and presented using Presentation software (Neurobehavioral Systems). Each trial began with a white fixation cross in the center of a black screen, after which two red squares (2.7° x 2.7°, 1.85° below fixation) appeared on the left and right sides of the screen, respectively. The squares appeared 16.4° apart horizontally and were always equidistant from the fixation cross, which remained at the center of the screen. The squares were presented at 13 different stimulus-onset asynchronies (SOAs): -96, -60, -48, -36, -24, -12, 0, 12, 24, 36, 48, 60, and 96, where negative SOAs indicate that the square on the left side of the screen appeared first and positive SOAs indicate that the square on the right side of the screen appeared first. The SOA of 0 indicates that both squares appeared simultaneously, but participants were not informed the squares may appear simultaneously. On each trial, participants were to make a judgment regarding whether the square on the left had appeared first or the square on the right had appeared first and indicate their responses using the ‘1’ and ‘2’ keys respectively. Participants pressed the ‘0’ key between each trial in order to begin the next trial. After a practice block of 26 trials, participants completed 260 experimental trials (10 blocks of 26 trials each) with a different random trial order presented to each participant.
2.2. Results

We provide results for the primary task (multiple-target visual search) first, followed by results for the control task (temporal-order judgment).

2.2.1. Data preparation and analyses

For all results, effect size and confidence intervals are reported (see Fritz, Morris, & Richler, 2012 calculation recommendations). Specifically, effect size was assessed using a modified calculation of Cohen’s $d$ (Cohen, 1962) recommended when the groups are similar in sample size but may have different standard deviations (Cohen, 1988; Keppel & Wickens, 2004). Resulting values were further adjusted to account for the small sample size and provide a more conservative estimate of the effect size, called $d_{\text{unbiased}}$ (Borenstein, Hedges, Higgins, & Rothstein, 2009), reported as $d_{\text{unb}}$. Like the standard Cohen’s $d$, $d_{\text{unbiased}}$ values of 0.8, 0.5, and 0.2 are generally representative of large, medium, and small effect sizes (Cohen, 1988). The 95% confidence intervals for effect sizes were calculated as recommended for small sample sizes and normally distributed data (Grissom & Kim, 2005; Hedges & Olkin, 1995).

We limited analyses to trials on which participants clicked the ‘DONE’ button before the time limit expired for two reasons: (1) to provide a conservative assessment of search performance (e.g., by excluding second-target errors related to incomplete searches) and (2) because the radiologists were significantly more likely to reach the time limit than the non-professionals (Radiologists: Mean=25.05%, SD=19.41%; Non-
professionals: \( \text{Mean}=2.08\%, \text{SD}=1.44\%; t(16)=3.76, p<0.001; d_{unb}=1.59\pm0.30 \). By restricting analyses to only self-terminated trials, we can be assured that we are focusing on comparable data across the two groups. Importantly, doing this did not skew the relative distribution of trial types across our two population groups: A 2 x 4 ANOVA on the number of trials excluded with Group (Radiologists vs. Non-professionals) as a between-subjects factor and Trial Type (single-target high salience trial, single-target low-salience trial, dual-target trial, and no-target trial) as a within-subjects factor revealed that there were no significant differences across the types of trials excluded for each group, \( F(3,17)=0.70, p=0.55 \).

### 2.2.2. Accuracy

#### 2.2.2.1. Overall accuracy

The groups did not differ in overall search accuracy (i.e., the percentage of all trials completed correctly — no misses and no false alarms; \( t(16)=0.01, p=0.99, d_{unb}<0.01\pm0.93 \)). Non-professionals produced a slightly higher false-alarm rate than radiologists (\( t(16)=2.17, p=0.046, d_{unb}=0.99\pm0.99 \)), but false alarms were quite rare in both groups (see Table 1 for specific values).

<table>
<thead>
<tr>
<th>Accuracy (%)</th>
<th>Radiologists</th>
<th>Non-professionals</th>
<th>Statistical Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>84.81 (5.95)</td>
<td>84.83 (4.18)</td>
<td>( t(16)=0.01, p=0.99 )</td>
</tr>
<tr>
<td>Single-target (high salience)</td>
<td>92.13 (7.91)</td>
<td>96.12 (3.29)</td>
<td>( t(16)=1.45, p=0.17 )</td>
</tr>
<tr>
<td>Single-target (low salience)</td>
<td>70.60 (13.53)</td>
<td>66.30 (9.87)</td>
<td>( t(16)=0.78, p=0.45 )</td>
</tr>
<tr>
<td>False alarm rate</td>
<td>0.83 (0.80)</td>
<td>1.73 (0.92)</td>
<td>( t(16)=2.17, p=0.046 )</td>
</tr>
</tbody>
</table>
2.2.2.2. Single-target accuracy

There were no differences between the groups in accuracy on single-target trials, whether the trials contained a high-salience target ($t(16)=1.45$, $p=0.17$, $d_{unb}=0.63\pm0.95$) or a low-salience target ($t(16)=0.78$, $p=0.45$, $d_{unb}=0.35\pm0.94$).

2.2.2.3. Relative second-target accuracy

The satisfaction-of-search (SOS) effect was examined by assessing relative accuracy rates for second targets. SOS scores were calculated by comparing accuracy for low-salience targets on single-target trials versus low-salience targets on dual-target trials in which the high-salience target was detected first (e.g., Cain & Mitroff, 2012). Non-professionals demonstrated a significant SOS effect ($t(9)=2.62$, $p=0.03$, $d_{unb}=0.69\pm0.91$), and radiologists showed a marginally significant SOS effect ($t(7)=2.33$, $p=0.05$, $d_{unb}=0.91\pm1.03$); there were no differences between the groups in SOS effects ($t(16)=0.52$, $p=0.61$, $d_{unb}=0.23\pm0.93$).

The above SOS calculation simplifies analyses by limiting SOS comparisons to only low-salience target performance. This is a reasonable approach when participants demonstrate a strong bias to find the high-salience target first (e.g., Cain & Mitroff, 2012), but fails to account for the majority of the data in the current study, as radiologists found the low-salience target first on nearly half of the dual-target trials (see “Target-detection order” below). As such, a different SOS calculation was employed and involved three steps: (1) low-salience target accuracy was compared between
dual-target trials in which the high-salience target was found first and single-target low-salience trials (same as above); (2) high-salience target accuracy was compared between dual-target trials in which the low-salience target was found first vs. in single-target high-salience trials; (3) these two calculations were averaged, each weighted by the number of contributing cases. (For example, if a participant found the high-salience target first on 60% of the trials part 1 of the equation was multiplied by 0.4 and part 2 was multiplied by 0.6.) Using this more inclusive calculation, an SOS effect was observed for both the non-professionals ($t(9)=3.10, p=0.01, d_{unb}=0.92±0.93$) and the radiologists ($t(7)=2.42, p=0.046, d_{unb}=0.81±1.02$); again, there were no group differences in SOS effects ($t(16)=0.21, p=0.84, d_{unb}=0.09±0.93$).

### 2.2.3. Response time

#### 2.2.3.1. Total trial time

Across all trial types, radiologists spent significantly longer per trial than the non-professionals ($t(16)=3.07, p<0.01, d_{unb}=1.36±1.04$) and demonstrated relatively higher response times in each trial type (see Table 2 for response times for each trial type).

<table>
<thead>
<tr>
<th>Response Time (seconds)</th>
<th>Radiologists</th>
<th>Non-professionals</th>
<th>Statistical Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>11.49 (1.59)</td>
<td>9.50 (1.17)</td>
<td>$t(16)=3.40, p&lt;0.01$</td>
</tr>
<tr>
<td>Single-target (high salience)</td>
<td>11.73 (1.70)</td>
<td>9.63 (1.21)</td>
<td>$t(16)=3.40, p&lt;0.01$</td>
</tr>
<tr>
<td>Single-target (low salience)</td>
<td>11.77 (1.59)</td>
<td>9.79 (1.27)</td>
<td>$t(16)=3.28, p&lt;0.01$</td>
</tr>
<tr>
<td>Dual-target</td>
<td>10.68 (1.65)</td>
<td>8.82 (0.94)</td>
<td>$t(16)=3.35, p&lt;0.01$</td>
</tr>
<tr>
<td>No-target</td>
<td>11.29 (1.87)</td>
<td>9.47 (1.32)</td>
<td>$t(16)=2.76, p=0.01$</td>
</tr>
</tbody>
</table>
2.2.3.2. Individual target types

More information can be gleaned by assessing the time required to find the individual high- and low-salience targets within dual-target trials. Non-professionals consistently detected high-salience targets more quickly than radiologists (See Table 3).

Table 3: The time taken in seconds (with standard deviations) before identifying each target type for each trial type. For dual-target trials on which the high-salience target was found first, the response time for low-salience targets is the time between the identification of the high-salience target and the low-salience target. For all other measures, the reported response time is the time taken to identify the target type from the onset of the trial.

<table>
<thead>
<tr>
<th>Trial type</th>
<th>Group</th>
<th>High-salience targets</th>
<th>Low-salience targets</th>
<th>Statistical comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Radiologists</td>
<td>4.80 (1.22)</td>
<td>6.43 (1.11)</td>
<td></td>
</tr>
<tr>
<td>Single-target trials</td>
<td>Non-professionals</td>
<td>3.24 (0.65)</td>
<td>5.52 (0.73)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Statistical comparison</td>
<td>t(16)=3.49, p&lt;0.01</td>
<td>d_{unb}=1.52±1.07</td>
<td>t(16)=2.10, p=0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>d_{unb}=0.93±0.98</td>
</tr>
<tr>
<td>Dual-target trials (When high-salience target was found first)</td>
<td>Radiologists</td>
<td>3.74 (1.09)</td>
<td>4.21 (0.78)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-professionals</td>
<td>2.66 (0.53)</td>
<td>4.06 (0.74)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Statistical comparison</td>
<td>t(16)=2.79, p=0.01</td>
<td>d_{unb}=1.21±1.02</td>
<td>t(16)=0.98, p=0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>d_{unb}=0.19±0.93</td>
</tr>
<tr>
<td>Dual-target trials (All dual-target trials)</td>
<td>Radiologists</td>
<td>5.25 (1.34)</td>
<td>6.35 (0.93)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-professionals</td>
<td>3.10 (0.53)</td>
<td>6.12 (0.90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Statistical comparison</td>
<td>t(16)=5.01, p&lt;0.01</td>
<td>d_{unb}=2.01±1.16</td>
<td>t(16)=0.52, p=0.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>d_{unb}=0.24±0.93</td>
</tr>
</tbody>
</table>

This pattern was evident in both single-target trials (t(16)=3.49, p<0.01, d_{unb}=1.52±1.07) and dual-target trials on which the high-salience target was detected first (t(16)=2.79, p=0.01, d_{unb}=1.21±1.02). This pattern was also seen when comparing response times
across all dual-target trials regardless of the target detection order ($t(16)=5.01$, $p<0.01$, $d_{unb}=2.01\pm1.16$).

Unlike the group differences observed in time to identify the high-salience targets, radiologists and non-professionals demonstrated similar response times for finding low-salience targets. There was a marginally significant difference between response times for low-salience targets on single-target trials ($t(16)=2.10$, $p=0.05$, $d_{unb}=0.93\pm0.98$). For dual-target trials on which the high-salience target was detected first, the two groups spent equivalent amounts of time searching for the low-salience target ($t(16)=0.98$, $p=0.34$, $d_{unb}=0.19\pm0.93$). There were too few trials on which the non-professionals detected the low-salience target first to allow for a meaningful comparison of response times on those trials; however, there was no difference between the groups in the time to detect the low-salience targets on all dual-target trials, regardless of detection order ($t(16)=0.52$, $p=0.61$, $d_{unb}=0.24\pm0.93$).

2.2.3.3. Target-detection order

Analyses of search performance on dual-target trials revealed striking group differences in search patterns. As mentioned above, in previous studies using the present paradigm with non-professional searchers (e.g., Cain et al., 2011; Cain & Mitroff, 2012; Dowd & Mitroff, 2013; Fleck et al., 2010), the high-salience target is typically found before the low-salience target in dual-target trials. For the current study, on dual-target trials on which both targets were detected, the non-professional group demonstrated
this expected bias, finding the high-salience target first significantly more frequently ($t(9)=8.43$, $p<0.01$, $d_{unb}=5.08\pm1.87$). Radiologists, however, did not demonstrate the same bias; they were just as likely to detect the high-salience target first as the low-salience target ($t(7)=1.55$, $p=0.16$, $d_{unb}=1.05\pm1.05$; See Figure 2).

![Figure 2: Percentage of dual-target trials wherein the high-salience or low-salience target was found first for radiologists and non-professionals. Data include dual-target trials on which both targets were detected. Error bars represent standard error of the mean.](image)

A 2 x 2 ANOVA with Group (Radiologists vs. Non-professionals) as a between-subjects factor and Target Type Detected First (high-salience vs. low-salience) as a within-subject factor revealed a highly significant interaction ($F(1,16)=36.38$, $p<0.01$): non-professionals
found the high-salience target first significantly more frequently than radiologists 
(t(16)=4.26, \( p<0.01 \), \( d_{unb}=1.92\pm1.14 \)).

2.2.4. Control task (temporal-order judgment) results

Accuracy for the temporal-order judgment task was calculated as the percentage of trials on which participants correctly reported whether the square on the left or the right appeared first. Trials with an SOA of 0 (squares appeared at exactly the same time) were excluded from the analyses, and a 2x12 ANOVA was run on the remaining data with Group (Radiologists and Non-professionals) as a between-subjects factor and SOA (-96, -60, -48, -36, -24, -12, 12, 24, 36, 48, 60, and 96) as a within-subjects factor. As expected, there was a main effect of SOA, with accuracy higher for longer SOAs (\( F(1,16)=43.03, p<0.001 \)). However, there was no main effect of Group, (\( F(1,16)=0.002, p=0.97 \); post-hoc t-tests comparing accuracy on each of the 12 SOAs: all \( p>0.20 \)), indicating that performance of Radiologists and Non-professionals was equivalent on the control task.

2.3. Discussion

Conducting accurate visual searches can be a particularly difficult task, and error rates are exacerbated when searches are complex. The presence of more than one target adds substantial complexity to the search process and often gives rise to the satisfaction-of-search (SOS) effect, wherein a second target is less likely to be identified if one target has already been detected in the display (see Berbaum, 2012 for a review). Visual
searches in careers such as radiology often contain multiple targets, and the radiological community is aware of the SOS phenomenon, and radiologists are trained to avoid this pitfall. However, SOS errors are believed to account for almost one-third of false-negative errors in some forms of radiological search (e.g., acute cervical spine trauma; Anbari & West, 1997). Like professional searchers, laypersons tend to fall victim to the SOS effect as well (Fleck et al., 2010), suggesting that errors associated with searching for multiple-targets are a generalizable occurrence seen in both experienced and inexperienced populations. Recent research has begun to elucidate the causes of SOS errors in both radiological (e.g., Berbaum, Franken, Caldwell, & Schartz, 2010) and non-professional (e.g., Cain & Mitroff, 2012; Fleck et al., 2010) groups; however, prior to the current study, the two populations had not been directly compared in terms of their performance on multiple-target searches.

As in a related study (Nodine & Krupinski, 1998), radiologists spent significantly longer searching but showed no advantage over non-professionals on standard measures of search accuracy. In addition, both groups demonstrated an SOS effect. Beyond replicating previous work, the fact that the radiologists showed no advantage in any basic measure of search accuracy despite their increased response times speaks to potential confounds. Namely, the lack of a speed/accuracy trade-off for overall performance suggests that other accuracy differences are unlikely to result from a simple speed/accuracy trade-off and/or increased “effort” within the radiological group.
Additionally, the results from our control task indicate that radiologists performed quite similarly to non-professionals on a temporal-order-judgment task, which was unrelated to search—further supporting that potential differences in motivation are unlikely to be driving the results.

More nuanced measures of performance measures revealed substantial differences between the non-professionals and radiologists and suggest that the radiology group is using a different search approach relative to non-professionals. Our paradigm presented target stimuli of high- and low-salience to simulate the variability in ease of target detection in medical images (e.g., the severity of a fracture can greatly determine the detectability, Berbaum, El-Khoury, & Ohashi, 2007). Non-professionals found the highly salient targets quickly and with priority for the low-salience items, implying they began searching by conducting a general survey of the display and immediately identified items that stood out. Additionally, as in prior studies (e.g., Cain et al., 2011; 2012; Dowd & Mitroff, 2013), non-professionals typically identified the high-salience target before the low-salience target on dual-target trials.

The radiology population, however, produced a different pattern of results, suggestive of a more methodical search process. On dual-target trials, radiologists found the low-salience target first on nearly half of the trials, which was significantly more frequently than non-professionals. If radiologists strategically employ a pre-determined scan path on each trial (i.e., start in the upper left and zig-zag through the display), they
would be equally likely to encounter either target type first, regardless of its relative salience. Post-experiment questionnaire responses support this notion: All 8 radiology participants reported using a standard spatial search method on each trial, and only 2 of 10 non-professional participants reported doing so.

While we cannot make definitive claims regarding the scan paths employed by the participants as we did not monitor eye movement, the timing of the mouse clicks on each target type over the course of a trial can serve as a proxy for where and when participants were searching (e.g., Clark, Fleck, & Mitroff, 2011). Non-professionals were significantly faster to identify high-salience targets than radiologists, but the groups showed no difference in the time taken to identify low-salience targets. Consistent with the idea that non-professionals begin by broadly surveying the array, this type of search would yield a rapid identification of highly salient items, which is not observed in the radiologists.

The current search data suggest that the radiologists executed a methodical search process that was blind to target salience—they appeared to evaluate potential targets without doing a first pass search for just the high-salience items. If a searcher is behaving methodically without necessarily worrying about time restrictions, such a search pattern would lead to a tendency to exceed the time limit (which the radiologists had). Likewise, such a search pattern would produce a more even percentage of finding
the high- and low-salience targets first (which was true for the radiologists), and no group difference in the time needed to find low-salience targets (which was also true). By comparing visual search performance between radiologists and non-professionals, we found substantial differences that are suggestive of different approaches to the task. Specifically, we found evidence for SOS errors from both experienced radiologists and from non-professional searchers; however there were stark contrasts between the groups in terms of their search speed and biases towards detecting various target types. These results serve to inform the nature of multiple-target search and begin to elucidate how a variety of behaviors may lead to the same kinds of errors; further research using eye-tracking technology could confirm differences in scan paths and search patterns. These findings suggest that similar surface-level patterns of errors do not necessarily imply similar underlying patterns of searching, and the causes of multiple-target search errors may not be identical across populations.

2.4. Acknowledgements

For helpful conversation, we thank Elise Darling, Emma Dowd, Stephen Adamo, Adam Biggs, and Joe Volosky. This work was partially supported by the Army Research Office (#54528LS) and partially through a subcontract with the Institute for Homeland Security Solutions, a research consortium sponsored by the Human Factors Division in the Department of Homeland Security (DHS). This material is based upon work supported by the DHS under Contract No. HSHQDC-08-C-00100. Any opinions,
findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the official policy or position of DHS or of the U.S. Government. The study is approved for public release.
3. Enhanced change detection performance reveals improved strategy use in avid action video game players

As our everyday lives become increasingly more complex with technological advancements, it becomes increasingly more necessary to understand how extensive experience with specific activities can affect cognitive and perceptual abilities. Recent findings have revealed that individuals with extensive action video game experience consistently demonstrate improved performance across a variety of visual and attentional tasks when compared to individuals who rarely play action video games. For example, compared to non-video game players (NVGPs), avid action video game players (VGPs) respond more rapidly (e.g., Castel, Pratt, & Drummond, 2005; Dye, Green, & Bavelier, 2009; Orosy-Filders & Allan, 1989; Yuji, 1996), have improved spatial abilities (e.g., Okagaki & Frensch, 1994; Quaiser-Pohl, Geiser, & Lehmann, 2006; Terlecki & Newcombe, 2005), have enhanced temporal abilities (e.g., 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 (e.g., Karle, Watter, & Shedden, 2010), and have enhanced eye–hand coordination (Griffith, Voloschin, & Gibb, 1983). Further, VGPs demonstrate improved “low-level” visual abilities, or bottom-up processing, as seen in increased visual acuity (Green & Bavelier, 2007) and contrast sensitivity (Caplovitz & Kastner, 2009; Li, Polat, Makous, & Bavelier, 2009), as well as improved “higher-level” visual
abilities, such as enhanced top-down attentional control (Chisholm, Hickey, Theeuwes, & Kingston, 2010; Hubert-Wallander, Green, & Bavelier, 2010). Here we define “low-level” and “bottom-up” improvements as performance benefits involving physical changes in basic visual abilities (e.g., contrast sensitivity; Li et al., 2009) and “higher-level” and “top-down” improvements as performance benefits involving changes to higher cognitive processes such as shifts in attentional allocation (e.g., Chisholm et al., 2010) or strategy use.

From the broad array of prior video game research, it appears that action video game exposure may heighten and hone attentional abilities (Hubert-Wallander et al., 2010), thus guiding and enhancing performance in visually demanding tasks. An important question that often arises in regard to these striking benefits is about causality — are VGPs better than NVGPs because they have engaged in extensive action video game play or are they better because individuals with a pre-disposition to heightened attentional and perceptual abilities may be more likely to play such fast-paced, action packed video games? This issue has been addressed head-on by training studies in which NVGPs were exposed to video games and their subsequent performance approached that of VGPs (e.g., Green & Bavelier, 2003). Many studies have shown that trained NVGPs do reveal enhanced performance, which suggests a causal role of video game playing (e.g., De Lisi & Cammarano, 1996; De Lisi & Wolford, 2002; Dorval & Pepin, 1986; Green & Bavelier, 2003, 2006a,b, 2007; McClurg & Chaille, 1987; Okagaki &
Frensch, 1994; however, see Boot, Kramer, Simons, Fabiani, & Gratton, 2008; Gagnon, 1985; Rosenberg, Landsittel, & Averch, 2005; Sims & Mayer, 2002 for lack of training effects; and Nelson & Strachan, 2009 for more nuanced training effects). The issue of causality explores an important mechanistic explanation of VGPs’ benefits, but regardless of the causal nature of such benefits, differences between VGPs and NVGPs have been reliably demonstrated. The aim of the current study is to address an equally critical mechanistic question: can video game players excel in visual tasks, at least in part, because of enhanced strategy use?

Despite the myriad of VGP benefits, it remains unknown how VGPs outperform NVGPs. There are two feasible hypotheses that are not mutually exclusive. The bottom-up hypothesis suggests that action video game exposure develops low-level differences that allow for 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). According to this hypothesis, VGPs may have an increased capacity to process visual information compared to NVGPs. Alternatively, the top-down hypothesis suggests that video game playing leads to the development of enhanced higher-level abilities such as attentional control (Chisholm et al., 2010) for generalized use across a variety of visually demanding tasks. For example, Chisholm et al. (2010) found that enhanced attentional control in VGPs can modulate the potentially negative effects of bottom-up attentional capture in spatial orienting. In line with this hypothesis, VGPs 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). Further, the top-down hypothesis suggests that VGPs may be better able to select (and adjust) their current strategies given the situation and their short- and long-term goals. Though strategy differences in long-term VGPs have yet to be explored fully, short-term exposure to video games has been found to influence speed/accuracy trade-off strategies in visual tasks (Nelson & Strachan, 2009). Here, we look to investigate how differences in strategies employed by long-term VGPs and NVGPs may relate to their improved abilities. While strong evidence exists in support of the bottom-up hypothesis (e.g., Green & Bavelier, 2007; Li et al., 2009), and some evidence supports the role of top-down attention control (e.g., Chisholm et al., 2010), we hypothesized that extensive action video game play may also work to develop enhanced top-down strategies in VGPs. The role of strategy has been proposed before — a prior research project used the classic Posner cuing paradigm to investigate the possibility of reduced attentional costs in VGPs (Greenfield, DeWinstanley, Kilpatrick, & Kaye, 1994). While there was mention of possible enhanced strategy use in VGPs, the data were ambiguous, as the paradigm provided no information about the process participants used to complete the task.

Here we use a change detection task to investigate whether improved use of top-down strategy can contribute to the benefits seen in video game players. Change
detection is a commonly used visual task in which participants attempt to identify a visual change between two scenes temporally separated by a disruption (see Simons & Rensink, 2005 for a recent review). Change detection provides a powerful tool for exploring issues of visual attention and perception since successfully noticing a visual change across a disruption requires forming, maintaining, and comparing visual representations (e.g., Mitroff, Simons, & Levin, 2004; Simons, 1996). These three necessary components of successful change detection tap into aspects of visual perception, attention, and memory, and each of these processes has been found to be enhanced in VGPs.

Change detection offers a nice tool for the current question since prior change detection studies have isolated various aspects of the detection process to examine the nature of visual processing. For example, change detection has been used to explore the contents of visual memory (e.g., Hollingworth & Henderson, 2004; Rensink, 2002; Simons, 1996) and the role of focused attention (e.g., Rensink, O’Regan, & Clark, 1997; Scholl, 2000). Further, failures to successfully detect a change have been used as evidence that 1) not all information is properly encoded (e.g., O’Regan & Noë, 2002), 2) even when information is encoded, it can be subsequently overwritten by new information (e.g., Beck & Levin, 2003; Levin, Simons, Angelone, & Chabris, 2002), and 3) even if information is not overwritten, two viable representations need not necessarily be compared to one another (e.g., Angelone, Levin, & Simons, 2003; Hollingworth, 2003;
Mitroff et al., 2004; Wang & Mitroff, 2009). Moreover, it has been argued that explicit processes may be necessary for the ultimate detection of a change (e.g., Mitroff & Simons, 2002; Mitroff, Simons, & Franconeri, 2002), which allows for us to focus the current questions at the level of perception with awareness.

A commonly used change detection paradigm is the “flicker” task (Rensink et al., 1997). In this task an image and a modified version of the image continuously alternate (with a blank display in between) until the participant finds the change between the two images. For the current goal of exploring how VGPs perform differently from NVGPs, we will employ a variant of the flicker task that allows for a step-by-step examination of the change detection process (Mitroff & Simons, 2002). This modified task has previously been used to look at group differences (young vs. older adults, Costello, Madden, Mitroff, & Whiting, 2010), making it especially compelling for the current goals. In this modified paradigm, individuated presentations of pre-change and post-change image pairs are presented one at a time and then the participants attempt to localize the change. The benefit of this design is that it slows down the change detection process and makes it possible to acquire multiple localization responses for each specific trial leading up the eventual detection (or non-detection) of a change (Mitroff & Simons, 2002).
3.1. Methods

3.1.1. Participants

35 male participants (VGP: N=15, mean age=19.93\(^1\), SD=1.69; NVGP: N=20, mean age=23.05, SD=5.87; \(t\)(32)=1.93, \(p=0.063\)) from the Duke University community received either course credit or $10 for a single 60-minute testing session. All participants had normal or corrected-to-normal vision.

Three additional participants (2 NVGPs, 1 VGP) incorrectly reported seeing a change on more than 25% of the no-change trials, and their data were excluded from all further analyses. Participants completed a video game questionnaire that assessed their experiences with several video game genres (e.g., first-person shooters, role-playing, puzzle) over several time frames and their responses were used to classify them as VGPs, NVGPs, or neither (participants that did not qualify as a VGP or NVGP were not included in this study). Those with extensive action video game playing experience were classified as VGPs. These participants played action video games (primarily first-person shooter games) for more than 6 h per week over the 6-month period prior to testing. Participants with no action video game experience and little to no experience with other video games in their lifetimes were classified as NVGPs. Prior to their testing session, many of the participants completed a condensed version of our video game

\(^{1}\) Represents mean age in years for 14 of the 15 VGP participants. As the result of a clerical error, one subject’s age information is missing.
questionnaire as part of a large test battery and we used their responses to selectively recruit VGP s and NVGP s while not revealing why they were being recruited. After the completion of the experiment, all participants completed our full questionnaire, and these responses were used to determine their status as VGP s and NVGP s. As in previous studies (e.g., Green & Bavelier, 2006b), female participants were not included since few meet the criteria for being classified as a VGP.

3.1.2. Apparatus and stimuli

The experiment was run in a dimly lit room on a Dell Dimension E520 with a 19-inch CRT monitor with a 1024 × 768 pixel resolution and a screen refresh rate of 60 Hz. Participants were seated at a viewing distance of approximately 57 cm without head restraint. Stimuli were presented and responses were collected with Matlab 7 software and the Psychophysics Toolbox (Brainard, 1997).

The stimulus set has been used in previous papers (Costello et al., 2010; Mitroff & Simons, 2002; Simons, Franconeri, & Reimer, 2000), and a detailed explanation of its construction can be found in Simons et al., 2000. The stimuli consisted of 64 photographs of natural scenes that subtended a visual angle of 18.97° × 12.71°. The stimuli were presented at the center of the screen and were surrounded by a black background that filled the remainder of the monitor. Each photograph was modified with Photoshop to either add or remove one item/ region to create an original and modified version of each image with only one change between the two. The average size of the change was 3.35%
of the total area of the image and the change sizes ranged from 0.43% to 14.46%. Across the 64 image pairs used in the current experiment, the change was equally distributed in each quadrant of the image (i.e., 25% of the changes occurred in the upper-left quadrant). Forty-eight of the image pairs were used for the change trials and the remaining 16 were used for the no-change trials. Only one image from each image pair was presented on the no-change trials, with half using the image that contained the modified item and half using the image that did not contain the modified item.

3.1.3. Procedures

Note that the experimental procedures were based closely on those of Mitroff and Simons (2002) and Costello et al. (2010). All participants were given written and oral instructions prior to the start of the experiment. They completed four practice trials with feedback, all of which contained a change. The images displayed during the practice trials were not used in the experimental trials. Participants were not informed prior to participation whether or not there would be a change in every trial.

Each trial began with a single white fixation cross presented on a black background for 500ms. Following the fixation, the first presentation cycle was presented. A single cycle consisted of the first image of the scene pair displayed for 250 ms, a blank gray screen replacing the image for 100 ms, the second image of the scene pair displayed for 250 ms, and finally, a blank gray screen that matched the size of the images that remained until response (see Figure 3).
Participants then used the computer mouse to make a localization response to indicate the location of the change between the two images. After the localization response, participants were to indicate their level of certainty with key-presses associated with the terms “Guess,” “Verify,” and “Saw.” If the mouse click was a complete guess, as participants had not seen any change, they were to indicate “Guess.” “Verify” was to be used if participants believed they saw a change but needed another presentation cycle to be certain. Participants were only to indicate “Saw” if they were confident they had seen the change and had clicked in the correct location. If the participant responded “Saw” the trial would end and they would move to the next trial. Otherwise, the entire
presentation cycle would repeat and participants were to make another set of mouse click and key-press responses. If the change was not detected after 15 total cycles, the next trial began. Thus for no-change trails, participants viewed all 15 cycles unless they falsely reported “Saw.” Participants were instructed to make their best localization guess at the end of each cycle (even if they did not see a change) and not to simply click the same location every time. Additionally, between cycles the mouse cursor was moved below the portion of the screen on which images were displayed to minimize haphazard responding.

The order of the 64 trials was uniquely randomized for each participant. Half of the change trials involved an addition within each presentation cycle (the changed item/region was present in the second image and not the first), and the other half involved a change deletion (the changed item/region was present in the first image and not the second). The changing region was defined as the smallest rectangular region that encompassed all changing pixels. For a few of the analyses below, we focused only on trials with a correct localization response, and for these, clicks within 30 pixels of the border of this rectangle were considered accurate. A 30-pixel window was employed since participants were making mouse click responses on a blank gray screen, and therefore, may not have clicked in exactly the right location, despite having successfully located the change. Since the trial only ended when participants indicated that they
“Saw” the change, this window did not increase the likelihood of a chance localization response influencing the data analyses.

3.2. Results

Overall performance was consistent with prior instantiations of this paradigm (e.g., Costello et al., 2010; Mitroff & Simons, 2002) with participants successfully finding 92.45% of the changes (SD=6.47%). False alarms on no-change trials (reports of “Saw”) were minimal and did not differ between VGPs (M=3.75%, SD=9.97%) and NVGPs (M=5.95%, SD=11.09%; t(33)=0.61, p=0.54). The “Verify” response was seldom used, and there was no significant difference in its use between VGPs (M=0.34 times per trial, SD=0.19) and NVGPs (M=0.33, SD=0.19; t(33)=0.11, p=0.92).

Several differences arose between VGPs and NVGPs and the primary findings are presented in Table 4. On change trials, VGPs took significantly fewer cycles to find the change than NVGPs (see Table 4 row A). While this result reflects that VGPs are better able to detect changes, four additional analyses offer insight into how VGPs were able to detect changes in fewer cycles than NVGPs, revealing that VGPs employed a broader search strategy in conducting their search.

The first three of these analyses were performed on mouse click response data from the no-change trials in which participants correctly viewed all 15 cycles without reporting “Saw” (no-change trials without a false alarm). First, VGPs’ successive localization responses were further apart than NVGPs’ (see Table 4 row B); from one
localization guess to the next, VGPs made mouse click responses that were a significantly greater distance from their previous mouse click response than did NVGPs.

Second, VGPs were more likely to make a localization guess in all four quadrants of the images (see Table 4 row C), suggesting they were more likely to search the entirety of the display. Third, VGPs were less likely than the NVGPs to perseverate in a single quadrant — on average, the NVGPs had a larger maximum number of localization responses within one quadrant (see Table 4 row D).

<table>
<thead>
<tr>
<th>Primary Analyses</th>
<th>Video game players (VGPs)</th>
<th>Non-video game players (NVGPs)</th>
<th>Statistical Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Number of cycles required to find change (change trials)</td>
<td>4.51 cycles (1.02)</td>
<td>5.32 cycles (1.05)</td>
<td>t(33) = 2.31</td>
</tr>
<tr>
<td>B Distance jumped from click to click (no-change trials)</td>
<td>7.77 deg (1.02 deg)</td>
<td>6.65 deg (1.73 deg)</td>
<td>t(33) = 2.20</td>
</tr>
<tr>
<td>C Tendency to cover all four quadrants (no-change trials)</td>
<td>14.00 trials (2.04)</td>
<td>10.90 trials (5.17)</td>
<td>t(33) = 2.19</td>
</tr>
<tr>
<td>D Maximum number of clicks per quadrant (no-change trials)</td>
<td>6.44 clicks (0.76)</td>
<td>7.51 clicks (1.89)</td>
<td>t(33) = 2.07</td>
</tr>
<tr>
<td>E Area covered in first-five clicks (change trials)</td>
<td>93.87 deg² [39.14 deg²]</td>
<td>78.71 deg² [41.27 deg²]</td>
<td>t(33) = 2.07</td>
</tr>
<tr>
<td>F Number of “undetected changes” (change trials)</td>
<td>8.20 trials (2.93)</td>
<td>8.19 trials (3.41)</td>
<td>t(33) = 0.01</td>
</tr>
</tbody>
</table>

A fourth analysis that was conducted on data from change trials in which the change was correctly detected, VGPs made localization responses across a significantly wider area of the image than the NVGPs (calculated as the area of the rectangle formed by the minimum and maximum clicks on the x- and y-axes, see Table 4 row E). To
ensure comparable data, we limited this analysis to the first 5 cycles on trials in which participants took 6 or more cycles to accurately locate a present change.

The above analyses show clear differences between VGPs and NVGPs; however, they are consistent with both the bottom-up and the top-down hypotheses. On one hand, and in line with the bottom-up hypothesis, VGPs may have been able to process more visual information during each fixation. This would explain why VGPs search more broadly by suggesting that during each cycle they were able to “eliminate” larger areas on the image as not containing a change, as they were able to process more visual information. On the other hand, and consistent with the top-down hypothesis, VGPs may have chosen to strategically employ a broader search strategy when looking for the changes. For these sorts of change detection tasks, participants are more likely to notice a change if they search the display broadly than if they perseverate in a given region. To differentiate between these competing explanations we examined what we termed “unrealized correct localizations” — trials on which participants successfully clicked on a change but failed to realize they had done so. Such occurrences were defined as the number of trials in which participants successfully made a localization mouse click on the change location, reported that the click represented a “Guess,” and then continued to search elsewhere for at least the next two cycles. On change trials, VGPs and NVGPs were not significantly different in their number of “unrealized correct localizations” (see Table 4 row F). There was no group difference in terms of when, within a trial, the
unrealized correct localizations occurred (VGP: $M = 3.54$ cycles into the trial, $SD = 3.68$; NVGP: $M = 3.95$, $SD = 1.97$ cycles, $t(33) = 0.45$, $p = 0.65$).

3.3. Discussion

Previous work has shown that VGPs outperform NVGPs on a variety of attentional and perceptual tasks and the current study sought to reveal how. It is informative to know that extensive action video game playing is associated with enhanced processing, but for this to become a viable research tool, we must understand what aspects of performance can be affected (e.g., Hubert-Wallander et al., 2010). Recent findings have offered support for “low-level” visual benefits in that VGPs show superior visual acuity and contrast sensitivity (e.g., Green & Bavelier, 2007; Li et al., 2009). Other work has suggested a possible “higher-level” benefit in the form of attentional control (e.g., Chisholm et al., 2010), but the role of top-down strategy had not been thoroughly investigated. Here we find additional evidence for higher-level claims in the specific form of strategy benefits and possibly reveal one manner in which VGPs can outperform NVGPs.

As predicted, and in line with prior work showing enhanced visual attention abilities in VGPs (e.g., Green & Bavelier, 2003, 2006a,b; Greenfield et al., 1994; Hubert-Wallander et al., 2010; West et al., 2008), VGPs performed better than NVGPs on our change detection task. When searching between two scenes for a change introduced during a disruption, VGPs required fewer exposures to the changing stimulus to detect
its presence. Interestingly, the current results do not support a prior study that revealed no differences in change detection performance between VGP s and NVGP s (Durlach, Kring, & Bowens, 2009). While it is not clear what led to these differing results, the current paradigm provides an arguably more sensitive means to tease apart how VGP s and NVGP s differed by slowing down the change detection process and collecting successive localization data leading up to the eventual detection (or miss) of a change.

The primary finding of the current experiment is the differences in search patterns between VGP s and NVGP s when searching for visual changes. VGP s exhibited broader search strategies, and in doing so, covered significantly more visual area. However, at first blush, these differences could be consistent with both a bottom-up and a top-down explanation: They could reveal that VGP s have better visual abilities and can encode more visual information on a given fixation (i.e., Green & Bavelier, 2007), or they could reveal enhanced change detection search strategies, such that VGP s make better use of the processed visual information, in that they choose to employ a broader search (i.e., Chisholm et al., 2010; Greenfield et al., 1994). In other words, VGP s make larger moves from one localization response to the next, for example, because either they can process more visual details than NVGP s from a given fixation (bottom-up) or because they choose to engage a broader endogenous search strategy (top-down). The results from our “unrealized correct localizations” analysis provide insight into this key explanatory issue. Unrealized correct localizations occurred when a participant made a
localization mouse click on the changed region of the image, but indicated the response to be a “Guess,” and then continued to search elsewhere. These occurrences reflect how often participants accidentally found the change and failed to realize that they had done so. If VGPs’ better visual abilities allowed them to take in more information on a given fixation, we would expect them to have a significantly lower number of unrealized correct localizations. If, however, a higher-level, broader search strategy is the driving force behind VGPs’ enhanced change detection performance, we would not expect a difference between VGPs and NVGPs on the number of unrealized correct localizations. The VGPs and the NVGPs did not differ in their rate of undetected changes (p=0.99), which suggests a strategy difference.

An assumption underlying several of the interpretations is that the mouse click data reveal the participants’ attentional allocation. While these mouseclicks do not provide a definite measure of attention, several factors suggest they are a useful proxy. First, participants reported post-experiment that they used the “Guess” mouseclicks to indicate where they planned to search on the next cycle of the trial. Second, the previous instantiation of this paradigm in Mitroff and Simons (2002) used a variety of permutations across experiments to reveal that participants used the mouseclicks to report their next locus of attention. Regardless, the different mouse click patterns between VGPs and NVGPs nevertheless reveal differences in strategy between the groups.
Both basic bottom-up visual abilities and top-down strategy choices are likely to be enhanced in VGPs, and the current results offer evidence for a top-down strategy contribution. This conclusion is consistent with other recent claims (e.g., Chisholm et al., 2010; Colzato et al., 2010) and reveals a generalizable VGP benefit that may not be tied to specific visual skills or paradigms. An exciting suggestion from the current study is that strategies obtained through extensive action video game playing may be a driving force in VGPs' benefits in visual attention tasks. The finding here that VGPs employ a broader search strategy for change detection raises the possibility that video game playing may introduce generalized training that can increase performance broadly. VGPs may learn to approach tasks more optimally and flexibly adjust their global strategies to meet the task at hand.

3.4. Acknowledgements

For helpful conversation, we thank Greg Appelbaum, Tym Blanchard, Matthew Cain, Sarah Donohue, Ricky Green, Joe Harris, Julia Schroeder, Grace Kohut, and Mike Schallmo. We also thank Melissa Bulkin and Ricky Green for assistance with data collection, David Madden for sharing programming code, and Wally Boot and an anonymous reviewer for insightful feedback. This research was partially supported by grants to SRM from the Army Research Office (#54528LS) and through a subcontract with the Institute for Homeland Security Solutions (IHSS). IHSS is a research consortium established to conduct applied research in the social and behavioral sciences. The
Human Factors Division is the Department of Homeland Security (DHS) sponsor for IHSS. More information about IHSS is available at http://www.ihssnc.org. This material is based upon work supported by the DHS under Contract No. HSHQDC-08-C-00100. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the official policy or position of DHS or of the U.S. Government. The study is approved for public release.
4. Improvement in visual search with practice: Mapping learning-related changes in neurocognitive stages of processing

Visual search, the process of detecting target items among distractors, is a vital cognitive ability central to many everyday human activities as well as to critical tasks such as detecting abnormalities in radiological images and screening airport luggage for contraband (see Wetter, 2013; Clark et al., 2012 for reviews). Successful search requires the execution of a cascade of fundamental cognitive processes, including sensory analysis of the scene, orienting of visual attention, working memory, target discrimination, and decision/response processes (see Nakayama & Martini, 2011; Eckstein, 2011 for reviews). These cognitive faculties are supported by various underlying neural mechanisms ranging from low-level feature analyses to higher-level, goal-driven decision processes (e.g., Corbetta & Shulman, 2002; Duncan & Humphreys, 1989; Treisman & Gelade, 1980).

Prior research has established that visual search can be improved through experience or practice (e.g., Sigman & Gilbert, 2000; Sireteanu & Rettenbach, 1995). Given the numerous neurocognitive stages involved in detecting, assessing, and responding to search stimuli, questions remain as to which processes are enhanced and in what relative combination. Several recent visual search studies have found amplitude changes in certain scalp-recorded event-related-potential (ERP) components after practice on complex conjunction search tasks (An et al., 2012; Hamame, Cosmelli,
Henriquez, & Aboitiz, 2011); however, the learning-related changes underlying training-induced behavioral improvement in the rapid, tightly timed, parallel processing involved in a feature-popout search tasks are unclear.

Here, we investigated changes across the entire stimulus-response processing cascade that underlie visual-search learning by leveraging the high temporal resolution of ERPs elicited to rapidly processed, feature-popout search targets. Participants completed a five-day behavioral practice protocol, and electrophysiological activity was recorded at the beginning and the end of the protocol to explore plasticity in the neural mechanisms underlying the expected reduction in behavioral response time.

To investigate the neural underpinnings of the anticipated improvement in behavioral performance (see Figure 4A), we assessed changes in four hallmark ERP components that reflect the cascade of cognitive processing stages from stimulus to response (see Figure 4B): (1) the posterior visual N1, a negative-polarity wave (latency ~150 ms) that reflects early sensory-evoked processing (Mangun & Hillyard, 1991); (2) the N2pc (negative-polarity posterior-contralateral, latency ~225 ms) associated with the shift of attention to a lateralized stimulus location (Luck & Hillyard, 1994a); (3) the SPCN (sustained posterior-contralateral negativity, latency 300-450 ms) or CDA (contralateral delay activity) that has been related to maintenance and manipulation of information in visual working memory (e.g., Ikkai, McCollough, & Vogel, 2010) and/or to cognitive processing required for target discrimination (Jolicoeur, Brisson, &
Robitaille, 2008); and (4) the *motor-related LRP component* (lateralized readiness potential) that reflects the initiation of a motor response (Coles, 1988).

Figure 4: Hypothetical model demonstrating potential changes in (A) behavior and (B) and ERP components. (A) Response time is expected to decrease after practice. B) Horizontal arrows indicate potential latency shifts in the N1, N2pc, and LRP components. Vertical arrows indicate potential amplitude changes in the P1/N1, N2pc, and SPCN components.

By comparing the neural activity associated with these cognitive processes before and after practice, we aimed to elucidate neural plasticity underlying the expected improvements in visual-search efficiency. Specifically, we assessed changes in the amplitude and/or latency of these ERP components to reveal how enhancements in sensory processing, attentional orienting, target discrimination, motor initiation, and/or motor execution contribute to improvement in visual search.

**4.1. Materials and Methods**

**4.1.1. Participants**

Nineteen healthy individuals with normal or corrected-to-normal visual acuity and normal color vision were recruited and provided informed consent. All procedures were approved by the Duke University Medical Center Institutional Review Board. All
individuals participated in a 5-day visual-search practice protocol for approximately 1-hour per day, over 5 consecutive days, beginning on a Monday and ending on the Friday of the same week. Prior to the start of the experiment on the first day, participants completed a brief (~5-minute) session to become acquainted with the task. Behavioral performance (accuracy and response time) was recorded on all 5 days, and scalp-recorded EEG was measured on the first and last days of the protocol. Participants provided informed consent and were compensated $15/hour.

Data from two participants were excluded from the analyses because of poor behavioral performance (accuracy percentages more than two standard deviations below the group mean). Data from four additional participants were excluded due to producing suboptimal EEG data (excessive eye or muscle artifacts) on one or both of the EEG sessions. Data from the remaining 13 participants (ages 18-35 years, 5 female) were included in all analyses.

4.1.2. Search paradigm

Stimuli were programmed and presented using the Presentation software suite (Neurobehavioral Systems, Albany, CA). During each of the 5 experimental sessions, participants completed a series of 14 blocks, each consisting of 150 trials and lasting approximately 4 minutes. Thus, each experimental session was comprised of 2100 trials and lasted approximately 56 minutes. Participants were seated, without head restraint, approximately 57 cm from the viewing monitor. A white fixation-cross was presented at
the center of the screen on a gray background and remained in place for the duration of each experimental block. Each trial consisted of a briefly presented (50 ms) circular array of 48 colored ellipses, of which 46 were blue, 1 was red, and 1 was green (see Figure 5), with each stimulus subtending a visual angle of 1.36° x 0.91°.

![Sample stimulus display](image)

**Figure 5: Sample stimulus display.** Blue ellipses are distractors, the green ellipse is the relevant color popout target, and the red ellipse is the irrelevant color popout non-target. Participants respond as to the orientation of the green target ellipse. In this example, a participant would respond by pressing the button corresponding to “horizontal.”

Thus, on each trial, there were two color-popout stimuli in the array, a green ellipse (the target) and a red ellipse (an irrelevant distractor). These green and red ellipses could
appear in one of 10 locations on the lower portion of the array on each trial and always appeared on opposite sides of each other. Participants were instructed to report the orientation of the green target ellipse as quickly and accurately as possible. Responses were made using the left and right fingers on a game controller to indicate “vertical” or “horizontal” orientations, respectively. Individual trials were separated by a stimulus-onset asynchrony that varied between 1300 and 1700 ms. Participants were instructed to maintain central fixation during the stimulus presentation to minimize eye movements and preserve visual stimulation consistency. The 50-ms duration of the presented stimuli was sufficiently brief so as not to allow for a saccade to the target.

The participants’ task was to find the green ellipse, shift their spatial attention to this target covertly (i.e., without making an eye movement) and discriminate its orientation (a larger vertical or larger horizontal aspect ratio) with a manual response. The design of the task enabled the extraction of several prototypical ERP markers of the cognitive processes involved in visual search. First, by structuring the search arrays with both a task-relevant target popout (green among blue distractors) and a task-irrelevant non-target popout (red among blue distractors), we were able to control for early sensory differences in the ERPs while also eliciting robust N2pc and SPCN components. Additionally, the manual responses were executed with the index fingers of the left and right hands (left for vertical targets and right for horizontal targets), thereby allowing
assessment of the lateralized readiness potential (LRP) associated with motor response initiation. Finally, the final response time was recorded for each trial.

4.1.3. Behavioral data acquisition and analyses

4.1.3.1. Behavioral analyses

Behavioral responses were considered accurate if the participant responded with the correct orientation (vertical or horizontal) of the target stimulus between 200 and 1000 ms following the onset of the array. Response time was recorded as the time between the onset of the stimulus array and the button press for each correctly reported trial. Within-subject differences in accuracy and response time over the course of practice were assessed using repeated-measures analyses of variance (ANOVAs). Additional two-tailed, paired t-tests were employed to compare accuracy and response time between the first (Session 1) and last (Session 5) experimental sessions.

4.1.3.2. EEG recording and analysis

On Sessions 1 and 5 of the experimental protocol, EEG data were recorded as participants performed the visual search task. EEG was recorded continuously using a custom extended-coverage elastic cap with 64 equally-spaced channels (Electro-Cap International, Eaton, OH), which covered the full head from slightly above the eyebrows to below the inion (Woldorff, et al., 2002). Impedances of all channels were adjusted to below 5 kΩ; EEG was amplified within a frequency band of 0.016-100 Hz and digitized at a sampling rate of 500 Hz per channel (SynAmps, Neuroscan, El Paso, TX). Eye
movements were monitored with vertical and horizontal EOG channels and a closed-
circuit zoom-lens camera, and participants were given verbal feedback to encourage
fixation on the central cross. Recordings took place in an electrically shielded, sound-
attenuated, dimly lit experimental chamber.

For each participant, EEG data were selectively averaged to yield ERPs for the
various conditions. All channels were re-referenced to the algebraic mean of the two
mastoid electrodes. A digital, non-causal, 9-point (18 ms) running average filter was
applied to the ERP averages to reduce signal frequencies greater than 56 Hz at our 500-
Hz sampling frequency. Artifact rejection was performed off-line by discarding epochs
of the EEG contaminated by eye movements or eye blinks (EOG), excessive muscle-
related potentials, drifts, or amplifier blocking. Artifact rejection thresholds were pre-set
to +/-120 uV for vertical eye channels and +/-90 uV for all other channels and applied
from -200 ms to 1000 ms around the presentation of the visual search array. Thresholds
were minimally adjusted for each participant to retain the most trials while eliminating
the above sources of contamination, and these artifact-rejection thresholds were then
applied via a computer algorithm that was blind to the specific trial types. These
parameters led to an average trial-rejection rate of 16.9%, yielding an average of 1765
usable trials per participant.

All changes in ERP components were assessed using two-tailed, paired t-tests to
compare amplitude and/or latency differences between Sessions 1 and 5. Activity in
parietal-occipital channels on the left (PO3, PO5, P3) and right (PO4, PO6, P4) sides of the head was analyzed for changes in the N1 sensory-evoked component. Two distinct regions of interest (ROIs) were assessed with respect to N2pc-related activity, a parietal region (superior ROI; channels PO3, P3, P5, PO4, P4, P6) and an occipito-temporal region (inferior ROI; channels PO3, PO5, O1, PO4, PO6, O2), per Hopf et al. (2000). Activity in these regions was also analyzed for the SPCN component (collapsed across the superior and inferior parietal ROIs). Finally, central-frontal channels (C5, CP5, C6, CP6) were used for assessment of the motor-related LRP component.

4.1.3.2.1. Visual N1

To examine early sensory processing, EEG activity from parietal-occipital channels was used to calculate the amplitude and latency of the visual N1 component (Mangun & Hillyard, 1991). For this analysis, trials were divided according to whether the target appeared on the left or right side of the screen, and activity in sites contralateral and ipsilateral to the target was assessed. Mean amplitude measures of the N1 for each participant were taken in a 25-ms latency window centered around the peak of the N1 (140-165 ms) observed in the across-subject grand-average ERP, and these values were compared between Sessions 1 and 5 to assess changes in basic sensory processing with practice. Peak latencies of the N1 were also measured for each participant and compared between Sessions 1 and 5 to assess for changes in speed of basic sensory processing.
4.1.3.2.2. N2pc

ERP difference waves reflecting activity associated with the attentional-shift-related N2pc component were derived for superior (parietal) and inferior (parietal-occipital-temporal) regions (Hopf et al., 2000), calculated as the difference between the activity in posterior electrodes contralateral minus ipsilateral to the relevant popout target stimulus (Luck & Hillyard, 1994a). The resulting N2pc difference waves were compared between Sessions 1 and 5. The latencies of the peaks of the N2pc components were compared between Sessions 1 and 5 to determine whether the attentional shift to the target occurred sooner, relative to stimulus onset, after practice. In addition, N2pc amplitudes were compared between Sessions 1 and 5 in the 40-ms latency window immediately surrounding the peak. Across both sessions, the average peak of the N2pc occurred at 224.6 ms (SD=19.5 ms) post-stimulus in the superior ROI and at 229.3 ms (SD=21.5 ms) in the inferior ROI. Thus, amplitudes were assessed by comparing the mean amplitudes of the N2pc activity in the 40-ms window immediately surrounding the average peak time for each ROI.

4.1.3.2.3. SPCN

The SPCN component, which occurs after the N2pc, was also computed as a contralateral-minus-ipsilateral difference wave using the same calculation employed for the N2pc, but collapsed across superior and inferior ROIs and examined in a later time window. The amplitudes of the SPCN component were analyzed in a broad time
window following the N2pc component (340-480 ms post-stimulus) and compared between Sessions 1 and 5 to assess changes in target-discrimination processing. Previous research has indicated that this component tends to become smaller for target discrimination processes that are easier relative to those that are harder or that require more working memory (Jolicoeur et al., 2008).

4.1.3.2.4. LRP

Activity associated with the motor-related LRP component was calculated as the voltage difference between electrodes over the motor cortices contralateral versus ipsilateral to the hand used to execute the response on each trial (i.e., left for vertical orientation, right for horizontal orientation). More specifically, activity on the side of the head ipsilateral to the response hand was subtracted from the activity on the side of the head contralateral to the response hand (electrode sites C3’ and C4’), and the resulting LRP difference waves were compared between Sessions 1 and 5 to assess whether participants began to prepare motor responses more quickly after practice. To capture onset latency specifically, rather than peak latency, we calculated the fractional peak latency – the time at which the ERP waveform reached 30% of its peak amplitude (Mordkoff & Gianaros, 2000). Additionally, we directly compared the latency of the measured peaks of the LRP components between Sessions 1 and 5. Lastly, we compared the amplitudes of the LRP components between the two sessions, assessed in a 40-ms window around the grand average peak of the component observed in each session.
4.2. Results

4.2.1. Behavioral results: Response time and accuracy

As expected, participants responded more quickly after practice. This improvement was reflected by a significant main effect of Session on response time ($F(4,12)=3.76, p=0.008$), with an average decrease of 81.1 ms and a significant difference between Sessions 1 and 5 ($t(12)=10.01, p<0.001$; see Figure 6A and Table 5A). There were no differences in accuracy across the five Sessions ($F(4,12)=0.16, p=0.96$) and no difference in accuracy in the direct comparison between Session 1 and Session 5 ($t(12)=1.01, p=0.33$; see Figure 6B and Table 5B); thus, participants maintained the same level of accuracy for the duration of practice. The fact that response time decreased significantly with no sacrifice in accuracy suggests that the improvement in response speed was not the result of a speed/accuracy tradeoff.

| Table 5: Means and standard deviations for (A) response time and (B) accuracy, before and after practice. |
|-----------------|-----------------|-----------------|-----------------|
|                 |     Session 1    |     Session 5    |     Statistics  |
| A. Response time|   546.7 ms (56.2)|   465.7 ms (50.8)|   $t(12)=10.01, p<0.001$ |
| B. Accuracy     |     88.2% (6.8%)|     90.2% (6.1%)|   $t(12)=1.01, p=0.33$   |
Figure 6: Behavioral results for (A) response time and (B) accuracy, before and after practice. Response time decreased significantly over the course of practice, but there was no significant change in accuracy.

4.2.2. Electrophysiological markers

4.2.2.1. Early visual sensory processing: N1 effects

Peak amplitudes of the N1 were significantly larger in Session 5 than in Session 1 both for sites contralateral \( t(12)=3.85, p=0.002 \) and ipsilateral \( t(12)=3.76, p=0.003 \) to the target (see Figure 7A and Table 6A). This overall amplitude increase of the N1 component suggests a generalized enhancement in sensory processing of the stimulus arrays with practice.

Table 6: (A) Means and standard deviations of peak amplitudes of the sensory-evoked N1 component, collapsed across left and right targets. (B) Means and standard deviations of latency of the peak of the N1 component in response to targets on the left and right sides of the display.

<table>
<thead>
<tr>
<th></th>
<th>Session 1</th>
<th>Session 5</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Amplitude</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contralateral to target</td>
<td>-4.53 µV (2.35)</td>
<td>-5.85 µV (3.02)</td>
<td>( t(12)=3.85, p=0.002 )</td>
</tr>
<tr>
<td>Ipsilateral to target</td>
<td>-4.28 µV (2.26)</td>
<td>-5.49 µV (2.83)</td>
<td>( t(12)=3.76, p=0.003 )</td>
</tr>
<tr>
<td><strong>B. Latency</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contralateral to target</td>
<td>154.7 ms (13.7)</td>
<td>154.8 ms (12.2)</td>
<td>( t(12)=0.07, p=0.95 )</td>
</tr>
<tr>
<td>Ipsilateral to target</td>
<td>152.2 ms (13.7)</td>
<td>153.7 ms (12.0)</td>
<td>( t(12)=0.94, p=0.36 )</td>
</tr>
</tbody>
</table>
Figure 7: (A) ERP traces of the sensory-evoked N1 component, collapsed across left and right targets, demonstrating an increase in amplitude after practice. (B) Distribution of N1-related activity over the scalp in response to targets on the left and right sides of the display.

The analyses also showed that the N1 peak amplitudes (latency window 140-165 ms) were significantly larger in sites contralateral to the target stimulus than in sites
ipsilateral to the target stimulus in both sessions ($t(12)=4.33$, $p<0.001$). Given that the
stimulus arrays were controlled for left-right physical stimulus differences (a color
popout on each side), these effects were likely related to the analysis including the early
part of the contralaterality of the N2pc, which began in the middle of the N1 latency
range (see below). There was also a non-significant trend for an interaction between
laterality and session ($F(1,24)=4.01$, $p=0.07$). No differences in the peak latency of the N1
component were observed between any of the conditions (see Table 6B).

4.2.2.2. Allocation of attention: N2pc peak latency and amplitude

After practice, the N2pc component peaked significantly sooner (by ~18 ms). This
shift was evident at both the superior ($t(12)=3.46$, $p=0.005$) and inferior ($t(12)=3.40,
p=0.005$) ROIs (see Figures 8A-C and Table 7B).

<table>
<thead>
<tr>
<th></th>
<th>Session 1</th>
<th>Session 5</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Amplitude</td>
<td>Superior ROI</td>
<td>-0.97 µV (0.48)</td>
<td>-1.26 µV (0.48)</td>
</tr>
<tr>
<td></td>
<td>Inferior ROI</td>
<td>-1.39 µV (0.59)</td>
<td>-1.52 µV (0.65)</td>
</tr>
<tr>
<td>B. Latency</td>
<td>Superior ROI</td>
<td>232.9 ms (20.9)</td>
<td>216.3 ms (14.3)</td>
</tr>
<tr>
<td></td>
<td>Inferior ROI</td>
<td>239.5 ms (22.8)</td>
<td>219.1 ms (14.8)</td>
</tr>
</tbody>
</table>

Table 7: (A) Means and standard deviations of the amplitudes of the N2pc component, collapsed across
left and right targets (contralateral versus ipsilateral to the target popout). (B) Means and standard
deviations of the latencies of the N2pc component, collapsed across left and right targets (contralateral
versus ipsilateral to the target popout).
There was also a significant increase in the amplitude of the peak of the N2pc component in the latency window immediately surrounding the average time of the peak at the superior ROI ($t(12)=4.23$, $p=0.001$) but not at the inferior ROI ($t(12)=1.13$, $p=0.28$; see Table 7A).

4.2.2.3. Working-memory and target-discrimination resources: SPCN amplitude

There was a significant decrease in the mean amplitude of the SPCN component after practice ($t(12)=4.05$, $p=0.002$; see Figures 8A-B and 8D and Table 8), suggesting that practice facilitated target-discrimination processes requiring retention and/or manipulation of information in visual short-term memory.

<table>
<thead>
<tr>
<th>Session 1</th>
<th>Session 5</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplitude</td>
<td>$-0.29 , \mu\text{V (0.9)}$</td>
<td>$0.03 , \mu\text{V (0.9)}$</td>
</tr>
</tbody>
</table>

4.2.2.4. Motor-response preparation: LRP onset latency

The onset of the LRP component was significantly earlier in Session 5 than in Session 1 ($t(12)=4.36$, $p=0.001$; see Figure 9 and Table 9), reflecting earlier initiation of the motor response after practice. This was further supported by the latency of the peak amplitudes of the LRP component also being significantly earlier in Session 5 than in Session 1 ($t(12)=3.67$, $p=0.003$). No difference was observed between the amplitude of the LRP component in Sessions 1 and 5.
Figure 8: (A) ERP traces of activity used to calculate the differences waves for deriving the N2pc and SPCN components (contralateral versus ipsilateral to the target popout), collapsed across superior and inferior sites. (B) Difference waves displaying N2pc and SPCN components at the superior and inferior ROIs. (C) Distribution of N2pc-related activity over the scalp for Sessions 1 and 5. (D) Distribution of SPCN-related activity over the scalp for Sessions 1 and 5.
Table 9: Means and standard deviations of the onset latencies of the LRP component.

<table>
<thead>
<tr>
<th></th>
<th>Session 1</th>
<th>Session 5</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Onset latency</td>
<td>353.7 ms (14.9)</td>
<td>314.3 ms (24.7)</td>
<td>t(12)=4.36, p=0.001</td>
</tr>
<tr>
<td>Peak latency</td>
<td>425.4 ms (15.7)</td>
<td>388.8 ms (33.1)</td>
<td>t(12)=3.67, p=0.003</td>
</tr>
</tbody>
</table>

4.2.3. Summary of results

The results from this study delineate learning and plasticity in key phases of the neurocognitive processing chain associated with behavioral improvements in visual search performance with practice. Over the course of five sessions of practice on a visual-search popout task, participants became significantly faster to detect and discriminate targets without sacrificing accuracy. The electrophysiological measures of brain activity showed that this behavioral performance improvement was accompanied by a significant increase in the amplitude of the sensory-evoked visual N1 ERP component, an increase in amplitude and shortening in latency of the attention-sensitive N2pc, a significant decrease in the amplitude of the SPCN reflecting reduced target-discrimination resources needs, a significant quickening of the onset latency of the LRP reflecting faster motor-response initiation, and a significant quickening of the time between the motor-response initiation and response time.
Figure 9: (A) ERP traces of activity used to calculate contralateral versus ipsilateral differences waves used for deriving the LRP component (contralateral versus ipsilateral to the hand used for the motor response). (B) Difference waves displaying the LRP component. (C) Distribution of LRP-related activity over the scalp for Sessions 1 and 5.
4.3. Discussion

We aimed to elucidate which stages of neurocognitive processing are improved with practice on a visual search popout task. We observed a robust improvement in response time (~81 ms) and identified several neural processes associated with this improvement. Prior studies investigating neural mechanisms underlying practice-related changes have employed visual search conjunction search tasks (An, et al., 2012; Hamame, et al., 2011); however, conjunction searches may require numerous fixations before attention is allocated to the target (Treisman & Gelade, 1980), making it difficult to infer the precise timing of brain responses related to search and learning.

In the present experiment, we employed a feature popout search in which the target element possesses a feature that is absent from all distractor elements and thus captures attention very quickly. In such popout searches, the N2pc component is easily observed and reflects rapid attentional selection of the target stimulus (Luck & Hillyard, 1994a) and/or suppression of distractors (e.g., Hickey, Di Lollo, & McDonald, 2009; Luck & Hillyard, 1994b). Additionally, popout search provides minimal trial-to-trial timing variability, allowing for high fidelity in the ERP comparisons. Moreover, there is a complex series of cognitive processing stages from stimulus input to behavioral response output that are required for visual search, and little is currently known about how plasticity in these various stages enable the behavioral improvements seen with
practice. Each of these processing stages and their pattern of practice effects are described below.

4.3.1. Initial sensory processing

We did not observe any change in the latency of the early sensory-evoked occipital N1 component, indicating that speeding of basic visual sensory processing did not contribute to the ultimate speeding of response time. We did, however, observe a significant increase in the amplitude of the N1 component after practice, suggesting enhancement of early sensory processing with practice. A prior study did not find a change in N1 amplitude in response to practice in a conjunction search task (Hamame, et al., 2011) and suggested that cortical reorganization at this early sensory processing stage may not have been necessary or beneficial for such a task, whereas it was for our parallel processing, feature-popout task. The larger N1s in the present study could also have been the result of directing more attention toward the incoming stimulus array after practice, as stronger attention to stimulus input has been shown to produce larger N1s (e.g., Luck et al., 1994a).

4.3.2. Allocation of attention

The N2pc is a parietal-occipital ERP component (latency 175-300 ms) that reflects a lateralized shifting and focusing of attention to a specified target item (Luck & Hillyard, 1994a). After practice on the current paradigm, we observed larger amplitudes of the N2pc, indicating enhanced attentional orienting. This result is consistent with the
previously mentioned practice studies with conjunction visual search (Hamame, et al., 2011; An, et al., 2012), which reported a larger N2pc after practice. Additionally, we observed a significant shortening in the latency of the peak of the N2pc component after practice, which was not observed in the prior studies. It is possible that this latency effect was observed here, but not in the prior studies using conjunction searches, because the tighter timing and minimal trial-to-trial variability of the responses in our popout search task enabled a more precise temporal measure of the attentional shifting process. Alternatively, this latency effect may reflect differential practice effects for popout searches versus conjunction searches. In either case, this latency effect suggests that it is possible for the rapid process of attentional allocation to a feature-popout stimulus to become even faster with practice.

Importantly, response time improved by ~81 ms after practice, and the N2pc shifted by only ~18 ms. Thus, it should be noted that the majority of processing speed improvement appears to occur later in the cognitive cascade, although a larger N1 and N2pc at an earlier latency could contribute to the acceleration of later processing. Lastly, because there was no latency shift in the early sensory-evoked N1 component with practice, the N2pc was the earliest point in the cascade of search mechanisms in which learning appeared to speed processing.
4.3.3. Working-memory and target-discrimination

The SPCN, or sustained posterior contralateral negativity, is a lateralized ERP component associated with cognitive processing that occurs following attentional allocation to a lateralized target stimulus (Jolicouer et al., 2008). The amplitude of the SPCN has been related to the cognitive resources required to complete the task, as more cognitively taxing tasks (e.g., those that involve high working memory loads or difficult target discrimination) tend to elicit an SPCN with a higher amplitude than do simple tasks (e.g., Eimer & Kiss, 2010). Unlike the N2pc, the SPCN does not typically show a peak, but rather displays sustained activity occurring for several hundred milliseconds while visual information is processed.

We observed a substantial decrease in the amplitude of the SPCN after practice. Participants searched the same displays and completed the same task over the course of the practice protocol, so the objective difficulty of the task itself was unchanged. Accordingly, the SPCN amplitude decrease may reflect learning for the vertical-horizontal discrimination task such that it became easier after practice, thus requiring less neural resources being devoted to performing this discrimination.

4.3.4. Motor-response preparation

The LRP, or lateralized readiness potential, is a well-characterized, centrally distributed, negative wave measured over the motor cortices (Coles, 1988). This component reflects preparation for motor activity, thus providing a high-temporal-
resolution marker of the initiation of voluntary movement. We observed a significant shortening in the onset latency of the LRP after practice (by ~39 ms), reflecting an improvement in the amount of time required before initiating the response execution.

4.3.5. Temporal changes in cognitive stages

The earliest temporal change after practice was observed in the N2pc (~18 ms earlier), reflecting a faster attentional shift to the target. An additional change in latency was apparent in the onset of the LRP, which occurred ~39 ms faster relative to stimulus onset; however, approximately half of this change is accounted for by the shift in the N2pc. Thus, earlier preparation for the motor response contributes approximately ~21 ms to the overall improvement in response time. Accordingly, the latency shifts observed in the N2pc and the LRP, together, account for ~39 of the ~81 ms change in response time, indicating that additional speeding of processing occurred between the initiation and execution of the motor response.

4.3.6. Relationship to learning theory

These findings can be interpreted within the context of the prominent “reverse hierarchy” model of perceptual learning (Ahissar & Hochstein, 2004). Under this framework, learning is a top-down, attention-guided process in which modifications begin at high-level visual areas and work backwards towards the sensory input level where there is better signal and less noise. As such, initial performance is limited by the resolution of higher visual cortical areas, while after training performance is limited by
the resolution at lower visual levels. In the context of the present study, visual search learning is partly reflected by the enhanced amplitude and shortened latency of the N2pc. This component has been associated with neural generators in the parietal and occipital-temporal cortices (Hopf et al., 2000; Robitaille et al., 2010) and may reflect processing that is functionally positioned at the intersection between the sensory cortices and the frontal-parietal attentional control network (Corbetta & Shulman, 2002). Given this position, one interpretation of some of the observed effects is that practice increases the salience of the relevant low-level features of the scene such that attention is more quickly deployed to the relevant target. This may, then, result in facilitated downstream processing (e.g., by enhancing the target-discrimination stage), which would then be reflected in the modulation of the longer-latency SPCN component. However, the reverse hierarchy learning theory can account only for the earlier stages of the present results as we also observed learning effects at the motor-related stages, such as the shortening of the time needed between motor initiation (i.e., LRP onset) and motor output (i.e., the RT).

4.3.7. Conclusions

Our primary goal was to determine which cognitive processes underlying visual search are enhanced with practice and to delineate their relative contributions to improved behavioral performance. In parallel with a marked decrease in response time with practice, we observed a number of changes in the underlying neural activity
associated with specific cognitive mechanisms – namely, enhanced early sensory processing to the visual search array, enhanced and earlier attention orienting to the target item, decreased need of resources required for target discrimination, more rapid initiation of motor-response preparation, and more rapid execution of the motor response following that initiation.

Visual search is a complex but critical cognitive function that requires a cascade of component processes to be carried out successfully and effectively. The present findings elucidate specific practice-induced changes in the component neurocognitive stages underlying visual search and offer a principled method for probing the neural mechanisms underlying learning in this essential cognitive ability.

4.4. Acknowledgements

This work was supported by NIH grant R01-MH060415 to M.G.W.
5. Context matters: The structure of task goals affects accuracy in multiple-target visual search

Numerous careers require individuals to conduct difficult visual searches; for example, radiologists search medical images for abnormalities, and airport security screeners search luggage for contraband. Accuracy for these tasks is critically important, as any errors could result in fatalities, and career searchers are trained to detect target items with as few errors as possible. Nevertheless, radiologists, airport security screeners, and other highly trained professional searchers still regularly miss targets. As such, a primary goal in applied visual search research is to identify the causes of search errors with the ultimate goal of improving accuracy and performance (Clark, Cain, & Mitroff, in press).

Visual searches conducted by professionals often present a number of significant complexities. One particular difficulty arises because search arrays can contain more than one target—a medical image could contain multiple abnormalities (e.g., a tumor and a fracture), and a suitcase X-ray could contain multiple banned items (e.g., a water bottle and a gun). Research in academic radiology has investigated the challenges associated with searching for multiple targets and identified a phenomenon known as “satisfaction of search” (SOS; Smith, 1967), the idea that observers tend to be less accurate in detecting a second target after having identified one target in a display (see Berbaum, 2012, for a review). The SOS phenomenon was originally believed to result
from an early termination of search, assuming that an observer was “satisfied” with the meaning of the display after the identification of one target and discontinued searching (Tuddenham, 1962). However, further research suggests that this is not the primary cause of SOS because observers do continue to search after detecting one target (e.g., Berbaum, Franken, & Dorfman, 1991). Instead, the decline in second-target accuracy may arise because of attentional disruptions related to the identification of the first target and the depletion of available cognitive resources (Cain & Mitroff, 2012), resulting in faulty decision-making (Berbaum, Franken, & Dorfman, 1998) or faulty pattern recognition (Samuel, Kundel, Nodine, & Toto, 1995).

Most investigations of SOS have used radiologists as participants and medical images as stimuli (Berbaum, 2012), but recent experimental work in cognitive psychology has used non-professional participants and precise manipulations of simplified stimuli (e.g., Fleck, Samei, & Mitroff, 2010) to understand the nature of multiple-target visual search more generally (e.g., Cain, Dunsmoor, LaBar, & Mitroff, 2011; Cain & Mitroff, 2012; Fleck, et al., 2010). Non-professional participants who search simplified displays demonstrate decrements in second-target accuracy paralleling those seen in radiology, revealing that SOS is a generalizable search phenomenon and not specific to the radiological community. Furthermore, multiple-target search paradigms can be a useful means for investigating the impacts of nuanced cognitive processes; contextual factors such as anticipatory anxiety (Cain, et al., 2011) and time pressure
(Fleck, et al., 2010) can have substantial effects on second-target accuracy without altering accuracy for single-target searches.

Exploring how multiple-target search accuracy can be improved is critical because most professional searches occur in settings where multiple targets are possible, and errors can have a tangible and direct impact on health and national security. The goal of the current study is to investigate whether the structure under which searchers complete their tasks can affect accuracy. Both radiologists and airport security screeners conduct series of searches as part of their jobs, but they do so under different constraints: Radiologists typically operate with a \textit{fixed objective} (e.g., assigned to assess 45 mammography images), while airport security screeners are scheduled to search for a \textit{fixed duration} (e.g., scheduled to serve as an X-ray screener at the passenger checkpoint for a 30-minute period).

Both radiologists and airport security screeners are trained to maximize accuracy and, in effect, should be attempting the same process—carefully examining each display for potentially harmful targets, regardless of the number of cases yet to be scanned or the amount of time left before the end of a shift. However, it is well known that the conceptual framework of a situation can dramatically alter behavior. For example, a substantially larger proportion of respondents are likely to support a medical program if presented in terms of the proportion of lives saved rather than proportion of lives lost, despite identical results between the conditions (e.g., Tversky & Kahneman, 1981).
Given that contextual factors (e.g., anticipatory anxiety and time pressure) can have negative effects on second-target accuracy in a multiple-target visual search (Cain, et al., 2011; Fleck, et al., 2010), we hypothesized that the framework under which an individual searches could also potentially alter performance. Specifically, we tested whether there are differences in accuracy when a search is completed within a task structure similar to radiology (searching with a fixed objective) versus airport security screening (searching for a fixed duration).

To address this question, we tested non-professional participants using a version of an established multiple-target search task with simplified stimuli that has reliably induced the SOS effect (e.g., Fleck, et al., 2010) and demonstrated sensitivity to environmental contexts (e.g., Cain & Mitroff, 2012; Clark, Cain, Adcock, & Mitroff, 2011). Professional and non-professional searchers tend to produce comparable patterns of multiple-target errors (Biggs, et al., 2013); however, it is important to account for potential differences in motivation between these groups in order to compare their search behavior. Undergraduate research participants may not be as concerned with their accuracy as radiologists and airport security screeners, for whom an error could have fatal consequences. Since assessing goal-relevant performance is only meaningful if individuals are truly attempting to attain the goal (Locke & Latham, 1990; Erez & Zidon, 1984), and monetary incentives offer a simple means to strengthen goal commitment (Locke, Latham, & Erez, 1988), we provided a performance-based monetary incentive to
increase the likelihood that the participants would genuinely attempt to achieve the instructed task goals. Related work using this motivational structure and the same multiple-target search task found enhanced accuracy in financially motivated versus non-motivated conditions (Clark, et al., 2011).

In the current experiment, we compared multiple-target search accuracy among participants searching with a fixed objective versus a fixed duration\(^1\). Two groups of participants completed an experimental search paradigm in which they accumulated points for accurate searching and were informed that the individual who achieved the “best” performance out of a set of 10 participants would receive an additional $50 in compensation. The paradigm was identical in each of the two conditions except for the framework of the participants’ task goal: In the Fixed Objective condition, participants were to achieve a specified number of points as quickly as possible; in the Fixed Duration condition, participants were to accumulate as many points as possible during a specified number of minutes. For the Fixed Objective condition, “best” was defined as the individual who achieved the specified points goal in the shortest number of minutes; for the Fixed Duration condition, “best” was defined as the individual who achieved the highest number of points in the specified time period. Importantly, the two conditions

---

\(^1\) The paradigm employed here is meant to approximate the nature of searches conducted by radiologists and airport security screeners, but key manipulations are necessarily altered. For example, the Fixed Objective structure is similar to radiological searches, but true radiological searches use a “Fixed Trials” structure, as immediate accuracy information is not feasible. A “Fixed Trials” condition would have substantially altered the strategy such that speed would be irrelevant.
were structured such that the optimal strategy in both was identical—to maximize one’s rate of point accumulation.

5.1. Methods

5.1.1. Participants

Forty undergraduate students were recruited from the Duke University community; 20 were randomly assigned to each condition (Fixed Objective: Mean age=20.15 years (SD=1.46), 17 female; Fixed Duration: Mean age=19.70 years (SD=1.34), 13 female). Participants provided informed consent and received $15 for their participation. Each participant had a 10% chance of earning an additional $50—the best performer from each of two consecutively recruited cohorts of 10 participants in each condition received the $50 bonus (i.e., 4 total bonuses were awarded, 2 for each condition).

Participants were not informed of their relative performance at the time of testing. After collecting and analyzing data from each set of ten participants, bonus recipients were contacted via email and invited back to the laboratory to collect payment. All other participants were notified via email that they had not received the bonus but thanked for their participation.

5.1.2. Apparatus

Stimuli were presented on a Dell Inspiron computer with a 20-inch CRT monitor and programmed in MATLAB (The MathWorks, Natick, MA) using the Psychophysics Toolbox (Version 3.0.8, Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007).
Participants were seated without head restraint at a viewing distance of approximately 57 cm from the screen and completed the experiment individually in a dimly lit room.

### 5.1.3. Design

Participants completed a modified version of a multiple-target visual search task that reliably reveals an SOS effect (e.g., Cain & Mitroff, 2012, Cain, et al., 2011, Fleck, et al., 2010; See Figure 10). Each trial contained 25 items, consisting of a short bar (0.9° long) and a long bar (1.3° long), each 0.3° wide, which approached one another perpendicularly to form ‘T’ shapes and pseudo-‘L’ shapes. Target ‘T’ shapes were defined as items in which a short bar approached a longer bar at its exact midpoint; the remaining items were considered distractor pseudo-‘L’s and were defined as items in which the short bar approached the longer bar at any point other than its exact midpoint. The shapes subtended a total area of 1.3° x 1.3° and were presented on a rendered grayscale “cloudy” background with a brightness range of 10–50% black. Distractor pseudo-‘L’ shapes were always between 28–66% black, and target ‘T’ shapes were presented in two visibility levels: high-salience targets (relatively dark; 66–70% black) and low-salience targets (relatively light; 28–40% black). The high-salience targets were easier to detect and distinguish from the background and distractor items compared to the low-salience targets. There were 0, 1, or 2 targets on each trial; single-target trials contained a target of either relatively low or high salience, and dual-target trials contained one high-salience target and one low-salience target. Each stimulus was
placed within a randomly selected cell of an invisible 8x7 grid with a total stimulus space of 25.4° x 19.1°.

![Sample Trial](image)

**Figure 10: Sample Trial.** Example display for a dual-target trial. This display contains one high-salience target ‘T’ (far right, middle) and one low-salience target ‘T’ (middle, far bottom).

Each trial began with a fixation cross appearing for 0.5 s at the center of the screen, after which the cross was replaced with the search array of 25 items. Participants used the mouse to click on each item they determined was a target and had the option to correct a misclick by clicking a ‘CLEAR’ button at the bottom of the screen. There was no time limit for individual trials; the experiments were self-paced, and participants clicked a ‘DONE’ button at the bottom of the screen to terminate each trial. Four types of trials
were employed throughout the experiment: no-target trials, single-target trials with a high-salience target, single-target trials with a low-salience target, and dual-target trials with one high-salience target and one low-salience target. Trial distribution was based upon Fleck, et al. (2010; Experiment 3) and pre-determined with rates of 20% no-target trials, 40% high-salience single-target trials, 16% low-salience single-target trials, and 16% dual-target trials. Exact trial-type rates varied slightly across individuals, as participants completed a different number of trials, depending on their accuracy and speed within the task constraints (see Sections 5.1.5 and 5.2.1).

5.1.4. Scoring and feedback

Participants received 1 point for every trial completed correctly (no misses, no false alarms) and lost 2 points for every trial in which an error was made (either a miss or a false alarm)\(^2\). Feedback was provided after each trial regarding the number of points gained or lost on the trial (See Figure 11). When an error was committed and points were lost, the type of error was printed on screen (e.g., “You missed a target” or “You clicked on a non-target”), and mistakes were highlighted in red. Two facts about cumulative performance were displayed after each trial: a running tally of the total number of points accumulated thus far and the number of minutes that had elapsed.

\(^2\) This scoring procedure was implemented based upon the trial-type distribution and pilot data. A 2-point penalty for incorrect trials was required to prevent participants from strategically terminating the trial immediately after finding only one target. Because there were more trials with 1 target than 2, without a penalty, the optimal strategy would be to quickly accumulate points on single-target trials only without searching for second targets.
Participants viewed the feedback screen between trials for as long as they liked and then pressed the spacebar to proceed to the next trial. Time spent viewing the feedback screen did not contribute to the participants’ total elapsed time on experiment.

![Sample Feedback Screen](image)

**Figure 11: Sample Feedback Screen.** Example feedback screen displayed following trial completion. On this example dual-target trial, the participant identified the high-salience target (denoted by the small circle) but missed the low-salience target (denoted by the large circle; red in experiment).

### 5.1.5. Framing of task by condition

Participants in the *Fixed Objective* condition were provided a goal of 230 total points and informed that the participant who reached 230 points in the fewest number of minutes, within a set of 10 participants, would receive an additional $50. Participants
in the Fixed Duration condition were given 70 minutes to complete the task and informed that the participant who accumulated the highest number of points in the 70-minute period, within a set of 10 participants, would receive an additional $50. The two frameworks were designed to be roughly equivalent, as pilot data suggested that the acquisition of 230 points required an average of 70 minutes. The total number of experimental trials varied by participant, as participants completed as many trials as were necessary to reach 230 points or as many trials as were necessary to reach 70 minutes. Some participants in the Fixed Objective condition did not reach the full 230-point goal. Participants were scheduled for a 90-minute session, and auxiliary activities (informed consent, practice trials, etc.) typically required 20 minutes, allowing 70 minutes for the experimental trials. If participants in the Fixed Objective condition reached 90 minutes of total participation, the experimenter entered the testing room and terminated the experiment, even if the participant had not yet accumulated 230 points. As pilot data suggested that 70 minutes was the average time required to accumulate 230 points, it was expected that some participants would not achieve the goal in the time allowed; their data were still included in the analyses.

Prior to the experimental session, each participant completed a brief practice session with an experimenter present. The practice sessions for each condition were shortened versions of the task framework they would complete in the experimental session. Participants in the Fixed Objective condition were to accumulate 23 points in the
shortest amount of time; participants in the Fixed Duration condition were to accumulate as many points as they could in a 7-minute period. Feedback provided during the practice session was identical to that provided during the experimental session. After completion of the practice session, participants confirmed they understood the task, the experimenter left the room, and participants began the experimental session.

5.2. Results

For all results, effect size and confidence intervals are reported (see Fritz, Morris & Richler, 2012 for calculation recommendations). Effect size was assessed using a modified calculation of Cohen’s $d$ (Cohen, 1962) recommended when the groups are similar in size but may have different standard deviations (Cohen, 1988; Keppel & Wickens, 2004), yielding one version of Hedge’s $g$ (Hedges, 1982). Like Cohen’s $d$, Hedge’s $g$ values of 0.8, 0.5, and 0.2 are generally representative of large, medium, and small effect sizes, respectively (Cohen, 1988). The 95% confidence intervals for effect sizes were calculated as recommended for normally distributed data and reasonable sample sizes (Grissom & Kim, 2005; Hedges & Olkin, 1995).

5.2.1. Comparability of experimental parameters between conditions

The two conditions were designed to broadly comparable based on pilot data. There was no significant difference in the total number of experimental trials completed between the conditions (Fixed Objective: Mean=342.95, $SD=54.15$; Fixed Duration: Mean=375.60, $SD=113.52$; $t(38)=1.16$, $p=0.25$; $g=0.37\pm0.63$). As specified by experimental
parameters, all participants in the *Fixed Duration* condition spent exactly 70 minutes on the task, and participants in the *Fixed Objective* condition spent an average of 61.69 minutes \((SD=9.32)\). Twelve of the 20 participants in the *Fixed Objective* condition reached the goal of 230 points before 70 minutes of time on task; for the remaining 8 participants, the experiment was terminated at 70 minutes, despite not having reached the goal.

5.2.2. Equivalent performance on basic measures

Participants in the *Fixed Objective* and *Fixed Duration* conditions demonstrated similar performance in terms of both single-target accuracy and response time. There were no significant differences between the conditions for accuracy on single-target trials for either high-salience targets (*Fixed Objective*: Mean=96.88%, \(SD=2.20%\); *Fixed Duration*: Mean=97.65%, \(SD=1.58%\); \(t(38)=1.27, p=0.21; g=0.40\pm0.63\)) or low-salience targets (*Fixed Objective*: Mean=68.88%, \(SD=13.71%\); *Fixed Duration*: Mean=73.31%, \(SD=8.06%\); \(t(38)=1.25, p=0.22; g=0.39\pm0.63\)). False alarm rates (percentage of trials on which any non-target was clicked) were very low and did not differ between conditions (*Fixed Objective*: Mean=0.82%, \(SD=0.63%\); *Fixed Duration*: Mean=0.89%, \(SD=0.68%\); \(t(38)=0.31, p=0.76; g=0.10\pm0.62\)). There were also no differences between the conditions in terms of response time across any trial type (See Table 10). Finally, the rate of point accumulation was equivalent between the conditions (reported as *points per minute*; *Fixed Objective*: Mean=3.44, \(SD=0.94\); *Fixed Duration*: Mean=3.65, \(SD=1.07\); \(t(38)=0.67, p=0.51; g=0.21\pm0.62\)).
Table 10: Response Times by Trial Type. Means (and standard deviations) in seconds for each trial type in the Fixed Objective and Fixed Duration conditions.

<table>
<thead>
<tr>
<th>Trial Type</th>
<th>High-Salience Single Target</th>
<th>Low-Salience Single Target</th>
<th>Dual Target</th>
<th>No Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Objective</td>
<td>11.36 (3.59)</td>
<td>11.81 (3.35)</td>
<td>9.01 (1.81)</td>
<td>13.10 (8.82)</td>
</tr>
<tr>
<td>Fixed Duration</td>
<td>12.71 (5.02)</td>
<td>13.44 (5.69)</td>
<td>9.98 (3.34)</td>
<td>14.43 (6.42)</td>
</tr>
<tr>
<td>Statistical Test</td>
<td>$t(38)=0.98, \ p=0.34$</td>
<td>$t(38)=1.11, \ p=0.28$</td>
<td>$t(38)=1.14, \ p=0.26$</td>
<td>$t(38)=0.55, \ p=0.59$</td>
</tr>
</tbody>
</table>

5.2.3. Dual-target accuracy and satisfaction of search

There were significant differences between the conditions for second-target accuracy (See Figure 12A). For dual-target trials on which the high-salience target was found first, participants in the Fixed Duration condition were significantly more accurate in finding the low-salience target as well (Fixed Objective: Mean=58.06%, SD=13.93%; Fixed Duration: Mean=69.88%, SD=8.96%; $t(38)=3.19, \ p=0.002; g=1.01\pm0.66$).

Satisfaction of search (SOS) is calculated as the difference in accuracy for low-salience targets between single-target trials and dual-target trials in which the high-salience target was found first (e.g., Cain & Mitroff, 2012). Participants in the Fixed Objective condition revealed a highly significant SOS effect (10.82%; $t(19)=5.19, \ p<0.001; g=0.78\pm0.64$), but participants in the Fixed Duration condition did not (3.44%; $t(19)=0.17, \ p=0.17; g=0.40\pm0.64$) (See Figure 12A). To assess the degree to which the SOS effect was modulated by the Fixed Objective versus Fixed Duration conditions, a 2x2 ANOVA was run on the low-salience target accuracy data with Condition (Fixed Objective vs. Fixed Objective) and Condition (Fixed Duration vs. Fixed Duration) as factors.
Duration) as a between-subjects factor and Trial Type (single-target trials vs. dual-target trials) as a within-subjects factor. There were main effects of both Condition

\( F(1,38) = 6.21, p = 0.02; \ g = 0.71 \pm 0.64 \) and Trial Type \( F(1,38) = 19.87, p < 0.001; \ g = 0.64 \pm 0.64 \) as well as a significant Condition x Trial Type interaction \( F(1,38) = 5.33, p = 0.026; \ g = 0.73 \pm 0.64 \), indicating that the SOS effect was larger in the Fixed Objective condition compared to the Fixed Duration condition (See Figure 12B).

Additional analyses reveal that both groups remained relatively consistent in their performance over the course of the experiment. To assess performance over time, the data were divided into quarters for each participant (i.e., separated into the first, second, third, and fourth 25% of trials; see Figure 13). A Quarter factor (First, Second, Third, Fourth) was added to the above analysis, and there was marginal, but non-significant, main effect of Quarter on accuracy \( F(3,38) = 2.415, p = 0.067 \); post-hoc tests demonstrated a general decline in accuracy toward the end of the experiment, with accuracy significantly lower in the Fourth quarter of trials compared to the first \( (t(38) = 2.61, p < 0.01) \). Most important for the current questions, there was no significant interaction between Condition and Quarter \( F(3,15) = 0.619, p = 0.60 \), indicating that the differences in accuracy between the Fixed Duration and Fixed Objective conditions remained constant over the course of the experiment.
Figure 12: (A) Accuracy rates for low-salience targets in the Fixed Objective and Fixed Duration conditions: Single-target trials vs. dual-target trials (provided the high-salience target was detected first). Error bars represent standard error of the mean. (B) SOS difference scores (difference between accuracy rates in Figure (A) in the Fixed Objective and Fixed Duration conditions. Error bars represent standard error of the mean.
Figure 13: Accuracy Rates over Time. Accuracy rates for low-salience targets in single- and dual-target trials in the Fixed Objective and Fixed Duration conditions during each quarter of trials over the course of the experiment.

5.3. Discussion and conclusions

Detecting a second target after having detected a first target in a display is a cognitively challenging task, and accuracy for additional targets tends to be uniquely sensitive to contextual influences that do not disrupt single-target searches. For example, accuracy impairments specific to second targets are observed under conditions such as anticipatory anxiety (Cain, et al., 2011) and time pressure (Fleck, et al., 2010). Here, we find that the even the mere structure of an observer’s search goals can affect accuracy in the same manner as stressful contexts, resulting only in differences specific to accuracy for second targets. There was a significant decrease in second-target accuracy for participants who were searching with a specified objective compared to those who were
searching for a specified duration, suggesting that the structure of an observer’s search goals affects his or her accuracy in detecting multiple targets.

Our participants were non-professionals who were motivated to search accurately with a performance-based monetary incentive. Participants were randomly assigned to one of two conditions and completed identical experimental paradigms; the only difference between the conditions was the framework in which they completed the task. Critically, the two frameworks called for employment of the same optimal strategy; whether attempting to achieve an objective in the shortest amount of time or to accomplish as much as possible in a specified period of time, searchers are attempting to maximize search efficiency in both conditions. However, humans are prone to irrationally conceptualize constructs such that objectively identical frameworks can dramatically alter decisions and behavior (Tversky & Kahneman, 1981).

Our analyses reveal that search performance, for the most part, was quite similar between the conditions, and there is no evidence that participants in the opposing frameworks were consciously employing different strategies or approaches to the task. Participants in the two conditions spent an equivalent amount of time assessing the search arrays and performed equally well on trials containing only one target. The only difference between the conditions was the likelihood with which participants found the additional targets on dual-target trials, with superior accuracy for multiple targets for the Fixed Duration condition. That is, the group of participants who were instructed to
accomplish as much as possible in a specified time period found second targets more deftly than those who were instructed to achieve a specified goal in the shortest number of minutes. Unlike those in the Fixed Duration condition, participants in the Fixed Objective condition produced a satisfaction-of-search effect, showing a substantial decline in accuracy for second targets.

All participants completed a modified version of a search task that typically elicits SOS (Fleck, et al., 2010, Experiment 3), and both groups completed this task under motivated conditions, a manipulation which can alleviate the SOS effect (Clark, et al., 2011). Interestingly, only those in the Fixed Duration condition showed the benefits associated with monetary incentives for this task; participants in the Fixed Objective condition performed similarly to non-motivated individuals in prior studies (e.g., Fleck, et al., 2010). Both groups were incentivized with a monetary reward, but performance on cognitive tasks can be modulated by the nature of the motivation (e.g., Callan & Schweighofer, 2008; Murayama & Kuhbandner, 2011). Motivation to avoid a punishment, for example, can be particularly stressful and promote anxiety (Davis & Whalen, 2001; Lang & Bradley, 2009), resulting in a decline in cognitive performance (Murty, LaBar, Hamilton, & Adcock, 2011). All of our participants were motivated to earn a reward, and there were no punishments to avoid (unless the scoring penalty for errors is considered a punishment in itself); however, it is possible that the Fixed Objective condition elicited more anxiety and stress than the Fixed Duration condition.
As mentioned, previous work has demonstrated that certain contexts tend to exacerbate errors specific to second targets in dual-target displays. Fleck, et al. (2010) found decreased accuracy under time pressure: observers committed significantly more errors when trials had a 15-second time limit than a 30-second time limit, despite the fact that participants rarely exceeded either time limit, and there were no differences between response times in the 15- versus 30-second conditions. Cain, et al. (2011) found that second-target errors were exacerbated when observers were searching under anticipatory anxiety. In some experimental blocks, participants were aware that they may receive a brief, uncomfortable shock to the wrist; in others, participants were aware that they may hear a neutral tone. In both cases, the shocks and tones occurred completely independent of performance, but second-target accuracy was significantly worse when anticipating the possibility of receiving an aversive shock than hearing a tone.

Both time pressure and anticipatory anxiety are potentially stressful contexts, and given that stressful motivation may inhibit cognitive performance (Davis & Whalen, 2001; Lang & Bradley, 2009), it is entirely possible the Fixed Objective framework examined here induced perceptions of time pressure and/or anxiety in the participants. Factually, both conditions imposed the same amount of time pressure, as speed is equally critical when attempting to achieve the most points over a specified duration or when attempting to achieve a set number of points in the shortest amount of time.
However, the framework of achieving a workload goal as quickly as possible might feel significantly more stressful than accomplishing as much as possible in a pre-determined amount of time. With the present data, we can only speculate about the role of the psychological and physiological states of the observer in these frameworks, but the strong similarities between our results and prior investigations of multiple-target search accuracy suggest that stress may be a common denominator.

Future work can speak to the underlying mechanism by which the task constraints in the current study influenced performance, however, regardless of the specific mechanism our data support the notion that second-target accuracy is improved when observers are searching for a certain period of time compared to when observers are searching to achieve an objective. This finding has direct implications for the structure of constraints for career searchers: Airport security screeners currently conduct searches for a pre-determined duration regardless of how many bags they search in that time; radiologists, on the other hand, are typically aware of a number of cases to be scanned until the job is complete. As second-target accuracy is a substantial problem in the radiological community, our data suggest that radiologists could benefit from a change in protocol. Rather than assigning a number of cases to each doctor, radiologists could be assigned to assess cases for a certain amount of time. This procedural modification could effectively increase second-target accuracy without decreasing efficiency as we find identical speeds for searching in the two frameworks.
5.4. Acknowledgements

For helpful conversation, we thank Elise Darling, Emma Dowd, Stephen Adamo, Adam Biggs, and Joe Volosky. This work was partially supported by the Army Research Office (#54528LS) and partially through a subcontract with the Institute for Homeland Security Solutions, a research consortium sponsored by the Resilient Systems Division in the Department of Homeland Security (DHS). This material is based upon work supported by the DHS under Contract No. HSHQDC-08-C-00100. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the official policy or position of DHS or of the U.S. Government. The study is approved for public release.
6. Effects of motivation, feedback, and time pressure on multiple-target search performance

A variety of situational factors can facilitate or diminish performance on cognitive tasks. For instance, a positive mood can enhance performance on basic decision-making tasks (Nadler et al., 2010), and a state of mental fatigue leads to a decrease in cognitive control (van der Linden, Frese, & Meijman, 2014). Experimental conditions associated with reward or punishment have similar effects (e.g., Savine & Braver, 2010) as do experimentally induced states of anxiety and stress (e.g., Berggren & Derakshan, 2012). The manner in which such factors influence cognition has important implications for the ideal structure of occupational settings, especially those in which optimal performance is critical.

Careers such as radiology and airport security call for high levels of accuracy in visual searches, and searches required in these professions involve several complications: 1) target prevalence is extremely rare, and 2) any given search could contain more than one possible target. The problems associated with low levels of target prevalence (e.g., Wolfe, Horowitz, & Kenner, 2005; Wolfe et al., 2007; Fleck & Mitroff, 2007) and the potential for multiple targets (e.g., Fleck, Samei, & Mitroff, 2010; Cain, Adamo, & Mitroff, 2013) have been studied extensively, and both rare- and multiple-target searches have shown sensitivity (both positive and negative) to contextual factors. Declines in accuracy associated with rare targets may be alleviated with a monetary
incentive (Navalpakkam, Koch, & Perona, 2009), but multiple-target search errors may be exacerbated with experimentally induced anxiety (Cain, Dunsmoor, LaBar, & Mitroff, 2010).

The increased error rates associated with searching for multiple targets within an array are termed subsequent search misses (SSMs; e.g., Adamo, Cain, & Mitroff, 2013), previously referred to as satisfaction of search errors. SSM rates are quantified as the difference in accuracy for displays containing only one target and for the second target identified in multiple-target displays. SSM errors are a common problem in radiology, and this effect has been demonstrated in non-professional settings as well (e.g., Fleck et al., 2010). Recent work has begun to untangle the causes of these errors (e.g., Cain, Adamo, & Mitroff, 2013), finding that they may emerge as a result of the depletion of cognitive resources expended during the identification of the first target found. Furthermore, accuracy rates for second targets tend to be uniquely sensitive to environmental influences such as anticipatory anxiety (Cain et al., 2011); in states of induced anxiety, accuracy for second targets suffers, while single-target accuracy remains unchanged.

A recent study (Clark, Cain, Adcock, & Mitroff, 2014) found substantial differences in relative second-target accuracy between two groups of individuals conducting searches within opposing frameworks. Both groups were provided a monetary incentive for accurate searching and performance-related feedback.
throughout the task; however, they were aiming for different goals: one group was attempting to perform as many accurate searches as possible in a set time period (fixed duration), while the other was attempting to achieve a set number of accurate searches in the shortest time period (fixed objective). The two frameworks call for identical strategies and impose the same amount of time pressure on observers; likewise, most measures of basic accuracy produced identical results, but there was a significant difference between the groups in SSM rates. Unlike participants in the fixed-duration condition, participants in the fixed-objective condition demonstrated substantial declines in accuracy for second targets.

These results raise several questions with regard to why SSM errors arise and how they can be alleviated. Why did participants in the fixed duration condition not produce SSM errors, which typically occur in similar paradigms? Two primary differences between the experiment employed by Clark et al. (2014) and similar studies are that, in Clark et al. (2014), 1) participants were provided a monetary incentive for good performance, and 2) feedback regarding errors was provided on screen throughout the task. Improved cognitive performance is generally observed with the motivation of a monetary incentive (e.g., Camerer & Hogarth, 1999) and when feedback is provided (Balzer, Doherty, & O’Connor, 1989). More specifically, each of these factors has been found to improve performance on another type of difficult visual search task, in which
targets occur extremely infrequently and are often missed (Navalpakkam, Koch, & Perona, 2009; Wolfe & Horowitz, 2007).

If motivation and feedback both facilitate improved performance, however, why were SSM errors still observed in the fixed-objective condition (Clark et al., 2014), which also provided a monetary incentive and feedback? One possibility, in line with the work relating SSM errors to anxiety (Cain et al., 2011), is that the framework of the fixed-objective condition (i.e., complete a specified number of accurate trials as quickly as possible) was subjectively experienced as more stressful by the participants. Additionally, prior work has demonstrated a relationship between time pressure and SSM errors (Fleck et al., 2010).

Motivation, feedback, and time pressure are all relevant factors for careers dependent on accurate searching, but it is not clear how each of these elements may contribute to multiple-target search accuracy. To dissociate the impacts of these contextual factors on search performance, the current series of experiments examined multiple-target search accuracy between groups of participants searching within each context and within the combinations of each of the contexts.

6.1. Methods

6.1.1. Participants

Eighty-two total individuals from the Duke University community participated in the experiment. Two participants were eliminated for poor performance (overall
accuracy more than two standard deviations below the mean), and replacement participants were run, yielding ten participants in each of eight conditions (46 female, mean age=21.9 years; SD=5.48). All participants provided informed consent and received $15 for their participation. In some conditions, participants had the opportunity to receive an additional $50 for superior performance.

6.1.2. Apparatus

Stimuli were presented on a Dell Inspiron computer with a 20-inch CRT monitor and programmed in MATLAB (The MathWorks, Natick, MA) using the Psychophysics Toolbox (Version 3.0.8, Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007). Participants were seated without head restraint at a viewing distance of approximately 57 cm from the screen and completed the experiment individually in a dimly lit room.

6.1.3. Multiple-target search task

All participants completed a multiple-target visual search task similar to that employed in prior studies (e.g., Clark et al., 2014; Cain and Mitroff, 2012; Cain et al., 2011; Fleck et al., 2010; See Figure 14). The base parameters employed in all eight conditions are described in this section; variations specific to each conditions are detailed in the sections below.
Figure 14: Sample array on a multiple-target search trial.

The experiment consisted of 255 total trials, and each trial contained 25 items, consisting of a short bar (0.9° long) and a long bar (1.3° long), each 0.3° wide, which approached one another perpendicularly to form ‘T’ shapes and pseudo-‘L’ shapes. Target ‘T’ shapes were defined as items in which a short bar approached a longer bar at its exact midpoint; the remaining items were considered distractor pseudo-‘L’ s and were defined as items in which the short bar approached the longer bar at any point other than its exact midpoint. The shapes subtended a total area of 1.3° x 1.3° and were presented on a rendered grayscale “cloudy” background with a brightness range of 10-50% black. Distractor pseudo-‘L’ shapes were always between 28 and 66% black, and
target ‘T’ shapes were presented in two visibility levels: high-salience targets (relatively dark; 66-70% black) and low-salience targets (relatively light; 28-40% black). The high-salience targets were easier to detect and distinguish from the background and distractor items compared to the low-salience targets. There were 0, 1, or 2 targets on each trial; single-target trials contained a target of either relatively low or high salience, and dual-target trials contained one high-salience target and one low-salience target.

Each stimulus was placed within a randomly selected cell of an invisible 8 x 7 grid with a total stimulus space of 25.4° x 19.1° (see Figure 14).

Each trial began with a white cross appearing for 0.5 s at the center of the screen, after which the cross was replaced with the search array of 25 items. Participants used a computer mouse to click on each item they determined was a target, and a small blue circle appeared on clicked targets. Participants had the option to correct a misclick by clicking a ‘CLEAR’ button at the bottom of the screen. To terminate a trial, participants clicked a ‘DONE’ button at the bottom of the screen upon concluding each search. Four types of trials were employed throughout the experiment: no-target trials, single-target trials with a high-salience target, single-target trials with a low-salience target, and dual-target trials with one high-salience target and one low-salience target. Trial distribution was based upon Fleck et al. (2010; Experiment 3): 20% no-target trials, 40% high-salience single-target trials, 16% low-salience single-target trials, and 16% dual-target trials.
6.1.4. Conditions

Participants were randomly assigned to complete one of eight versions of the multiple-target search task. Each version employed the paradigm described above with the presence or absence of three factors of interest (motivation, feedback, and time pressure), yielding eight conditions (See Table 11):

<table>
<thead>
<tr>
<th>Condition</th>
<th>Motivation</th>
<th>Feedback</th>
<th>Time pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>B</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>C</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>D</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>E</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>F</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>G</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>H</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

6.1.4.1. Motivation

Per Clark et al. (2014), participants in conditions that included the motivation factor (Conditions E, F, G, & H) were informed that within their group of 10 participants, an additional $50 would be awarded to the individual who performed “best.” Quality of performance was quantified as the percentage of total trials completed correctly; that is, the percentage of trials on which any and all targets were identified (no misses), and no distractors were incorrectly marked as a target (no false alarms). After all 10 participants completed each condition (typically within one week of participation for any individual), the best performer was notified via email and invited to return to
the lab for receipt of payment. All other participants were notified that they had not achieved the $50 bonus and thanked for their participation.

Participants in conditions that did not include the motivation factor were not informed of the potential for a monetary performance-based reward and were simply compensated $15 for their participation. (Participants in motivated conditions also received $15 upon participation.)

6.1.4.2. Feedback

Per Clark et al. (2014), participants in conditions that included the feedback factor (Conditions C, D, G, & H) were provided on-screen feedback regarding errors after each trial; the feedback appeared when participants clicked the ‘DONE’ button. For correct trials (i.e., those on which the participant did not miss a target or false alarm), the words “Successful trial!” were displayed. For trials on which an error was committed, the type of error was printed on screen (e.g., “You missed a target” or “You clicked on a non-target”), and mistakes were highlighted in red. After viewing the feedback screen, participants pressed the spacebar to proceed to the next trial.

Participants in conditions that did not include the feedback factor were not provided on-screen feedback regarding their performance during the experiment.

6.1.4.3. Time pressure

Participants in conditions that included the time pressure factor (Conditions B, D, F, & G) were informed that each trial had a time limit of 15 seconds. They were to aim to
complete the search on each trial and click the ‘DONE’ button within the time limit. If the ‘DONE’ button was not clicked before 15 seconds after the start of the trial, no further responses were accepted and a message was displayed on screen, “You have exceeded the time limit for this trial. Please try to complete your search within the time limit.”

Participants in conditions without the time pressure factor were not informed of any time constraints and able to search each display for as long as they wished, before clicking the ‘DONE’ button to terminate the trial.

6.2. Results

6.2.1. Data filtering

Prior to analysis, individual trials were excluded if participants in conditions including the time pressure factor failed to click the ‘DONE’ button prior to the time limit. Additionally, in all conditions, any trial on which participants used the ‘CLEAR’ button to correct a misclick were excluded. Less than 0.05% of all trials were excluded for one or both of the above reasons, and the remainder of the trials were submitted for analysis.

6.2.2. Accuracy and response time results for each condition

The following accuracy measures were computed for each condition:

1) “Overall”: Percentage of trials completed correctly (no misses, no false alarms)

2) “High-salience single target”: Accuracy percentage for high-salience targets on single-target trials
3) “Low-salience single target”: Accuracy percentage for low-salience targets on single-target trials

4) “Low-salience dual target”: Accuracy percentage of low-salience targets on dual-target trials on which the high-salience target was identified first.

5) “SSM”: Difference between “low-salience single target” and “low-salience dual target” accuracy percentages. Subsequent search misses (SSMs) represent the relative decline in accuracy for the second target identified in dual-target trials compared to accuracy for the same types of targets in single-target trials. For each condition, significance of SSM scores was assessed using paired t-tests between the “low single” and “low dual” measures.

6) “False alarm”: Percentage of trials on which participants identified a non-target item as a target.

In addition to the above measures of accuracy, total trial time (“Response time”) was calculated for each condition, that is, the number of seconds spent searching before the participant terminated the trial by clicking the ‘DONE’ button.

Means and standard deviations for each measure in each condition are presented in Table 12. ‘X’ marks in the columns labeled ‘M,’ ‘F,’ and ‘T’ indicate the presence of the motivation, feedback, and/or time pressure factors, respectively, in each condition.

---

1 Analyses for dual-target trials were restricted to the trials on which the high-salience target was identified first, per Cain and Mitroff (2013).
Table 12: Means (and standard deviations) for accuracy for each relevant target type and trial type, SSM difference scores, false alarms, and response time (RT), by condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>M</th>
<th>F</th>
<th>T</th>
<th>Overall</th>
<th>High single</th>
<th>Low single</th>
<th>Low dual</th>
<th>SSM</th>
<th>False alarms</th>
<th>RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>84.43</td>
<td>96.89</td>
<td>65.83</td>
<td>51.24</td>
<td>14.72</td>
<td>1.49</td>
<td>11.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.82)</td>
<td>(1.81)</td>
<td>(21.10)</td>
<td>(24.71)</td>
<td>(10.62)</td>
<td>(1.41)</td>
<td>(4.08)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>81.65</td>
<td>95.66</td>
<td>72.50</td>
<td>55.60</td>
<td>16.90</td>
<td>2.20</td>
<td>9.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.56)</td>
<td>(3.42)</td>
<td>(12.42)</td>
<td>(10.57)</td>
<td>(11.84)</td>
<td>(3.60)</td>
<td>(0.56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>82.27</td>
<td>94.24</td>
<td>61.00</td>
<td>48.91</td>
<td>12.09</td>
<td>1.49</td>
<td>9.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.14)</td>
<td>(3.43)</td>
<td>(17.92)</td>
<td>(21.61)</td>
<td>(12.40)</td>
<td>(1.43)</td>
<td>(2.20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>83.48</td>
<td>94.80</td>
<td>61.50</td>
<td>53.96</td>
<td>11.95</td>
<td>1.25</td>
<td>8.94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.18)</td>
<td>(3.89)</td>
<td>(15.64)</td>
<td>(13.17)</td>
<td>(13.79)</td>
<td>(0.94)</td>
<td>(2.03)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>86.78</td>
<td>96.32</td>
<td>71.00</td>
<td>60.52</td>
<td>10.48</td>
<td>2.24</td>
<td>12.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(11.21)</td>
<td>(4.47)</td>
<td>(27.49)</td>
<td>(27.14)</td>
<td>(12.61)</td>
<td>(4.35)</td>
<td>(5.41)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>85.69</td>
<td>93.33</td>
<td>73.05</td>
<td>61.63</td>
<td>10.99</td>
<td>1.53</td>
<td>11.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.39)</td>
<td>(5.89)</td>
<td>(10.37)</td>
<td>(15.92)</td>
<td>(12.90)</td>
<td>(1.02)</td>
<td>(2.20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>90.04</td>
<td>96.40</td>
<td>83.00</td>
<td>67.51</td>
<td>15.49</td>
<td>1.06</td>
<td>16.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.12)</td>
<td>(4.07)</td>
<td>(12.52)</td>
<td>(18.53)</td>
<td>(14.60)</td>
<td>(1.01)</td>
<td>(6.50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>90.35</td>
<td>98.08</td>
<td>77.25</td>
<td>69.86</td>
<td>7.39</td>
<td>0.75</td>
<td>10.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.53)</td>
<td>(2.30)</td>
<td>(13.61)</td>
<td>(14.65)</td>
<td>(1.20)</td>
<td>(0.63)</td>
<td>(1.68)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All conditions elicited significant SSM errors (aside from Conditions D and E, in which the effect was marginally significant.

**6.2.3. Overall analyses**

ANOVA was employed to reveal global effects of each of the three factors (motivation, feedback, and time limit) on all measures of accuracy. For the following
overall analyses, conditions were collapsed according to the presence or absence of each factor.

6.2.3.1. Global impact of motivation

There was a main effect of motivation on overall accuracy ($F(1,7)=10.50, p<0.01$) and response time ($F(1,7)=12.67, p<0.01$). Additionally, there were significant main effects of motivation on both single-target ($F(1,7)=7.21, p<0.01$) and dual-target trials ($F(1,7)=7.20, p<0.01$) but no influence on high-salience target accuracy. Motivation did not have a significant overall impact on SSM errors or false alarms (all $p$’s > 0.4).

6.2.3.2. Global impact of feedback

Analyses revealed no main effect of feedback across all conditions and measures. However, there was a significant interaction between motivation and feedback on single-target-trial accuracy for both high-salience ($F(1,7)=5.50, p=0.02$) and low-salience ($F(1,7)=4.87, p=0.03$) targets.

6.2.3.3. Global impact of time pressure

There was a main effect of time limit only on overall response time ($F(1,7)=10.18, p<0.01$) and no significant interactions between time limit and either motivation or feedback.
6.2.4. Detailed analyses

6.2.4.1. Time pressure does not affect accuracy

Analyses in Section 6.2.3.3 revealed no main effect of time pressure on accuracy with all conditions collapsed. Given the effects of and interactions between other factors, however, follow-up t-tests were performed between each pair of conditions, matched according to the inclusion or exclusion of the motivation and/or feedback factors.

Paired t-tests revealed no differences in accuracy between any matched conditions in which the only difference was the presence or absence of a 15-second time limit. There were no differences between Conditions A and B (no motivation, no feedback), but a trend for slightly longer response times in Condition A (no time pressure), (t(18)=2.03, p=0.057). There were also no differences between Conditions C and D (no motivation, yes feedback) and no differences between Conditions E and F (yes motivation, no feedback). Comparing Conditions G and H (yes motivation, yes feedback), revealed only a response-time difference.

Because time pressure did not vary significantly with any measure of accuracy, and the analyses below produce nearly identical results when assessing matched groups with and without time pressure, data from the eight conditions were collapsed to exclude the time pressure factor for the remaining comparative analyses, yielding twenty participants in each condition (See Table 13).
Table 13: Means (and standard deviations) for all measures, collapsed across the time pressure factor.

<table>
<thead>
<tr>
<th>Condition</th>
<th>M</th>
<th>F</th>
<th>Overall</th>
<th>High single</th>
<th>Low single</th>
<th>Low dual</th>
<th>SSM</th>
<th>False alarms</th>
<th>Response time</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/B</td>
<td>83.04 (6.76)</td>
<td>96.23 (2.70)</td>
<td>69.25 (16.49)</td>
<td>53.44 (17.74)</td>
<td>15.81 (11.01)</td>
<td>1.84 (1.54)</td>
<td>10.46 (3.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C/D</td>
<td>X 82.81 (6.56)</td>
<td>94.52 (3.58)</td>
<td>61.25 (16.37)</td>
<td>51.44 (17.61)</td>
<td>9.81 (14.69)</td>
<td>1.37 (3.47)</td>
<td>9.26 (2.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E/F</td>
<td>X 86.24 (8.90)</td>
<td>94.96 (5.13)</td>
<td>71.13 (20.46)</td>
<td>60.39 (21.55)</td>
<td>10.74 (15.59)</td>
<td>1.88 (2.36)</td>
<td>12.23 (4.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G/H</td>
<td>X X 90.20 (11.21)</td>
<td>97.24 (3.33)</td>
<td>80.13 (13.07)</td>
<td>68.69 (16.30)</td>
<td>11.44 (13.74)</td>
<td>0.90 (0.27)</td>
<td>13.47 (5.54)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6.2.4.2. The combination of motivation and feedback improves accuracy

To assess individual contributions of motivation and feedback and control for interaction effects, paired $t$-tests were employed to compare accuracy measures for matched conditions with and without motivation and/or feedback. Analyses reveal no significant effect of motivation alone. For the conditions that did not include feedback, there were no significant differences between Conditions A/B (no motivation) and Conditions E/F (yes motivation).

Significant differences in accuracy are only apparent between motivated and non-motivated conditions in the presence of feedback. There were significant differences between Conditions C/D (no motivation) and G/H (yes motivation), which both included feedback. Participants incentivized with a performance-based monetary reward were significantly more accurate overall ($t(38)=3.18$, $p<0.01$) and more accurate for each type of target and trial, including, high-salience single-target trials ($t(38)=2.49$, $p<0.05$).
135

\( p=0.017 \), low-salience single-target trials \((t(38)=4.30, p<0.001)\), and dual-target trials \((t(38)=3.22, p=0.003)\).

6.3. Discussion

In this series of experiments, we examined the effects of motivation, feedback, and time pressure on multiple-target search performance. Individual groups of participants completed one of eight variations of a multiple-target search task that included or excluded each factor: 1) the potential to receive a performance-based monetary reward (motivation), 2) information regarding errors after each trial (feedback), and 3) a 15-second time limit for each trial (time pressure).

Relative declines in accuracy for second targets in multiple-target searches (SSMs) are a common problem in radiology (e.g., Berbaum, 2012) and occur in non-professional populations as well (e.g., Fleck et al., 2010). Furthermore, SSM rates tend to be especially sensitive to contextual factors (e.g., task framing; Clark et al., 2014; anticipatory anxiety; Cain et al., 2010). In such cases, most measures of accuracy remain unchanged by the manipulation; there are no differences in accuracy for targets on single-target trials. The critical effect is typically observed in the relative decline in accuracy for low-salience targets on dual-target trials, or SSM rate. In this series of experiments, however, SSM rates were at least marginally significant in every condition, and accuracy differences were observed for both single- and dual-target trials.
Contrary to the conclusions discussed by Fleck et al. (2010), there was no effect of time pressure on any measure of accuracy, and SSMs occurred in conditions with and without time pressure alike. These differences in results can be explained by a discrepancy in the method for calculating SSM errors between the two studies: in Fleck et al. (2010), SSM rates were calculated as the difference between accuracy for low-salience targets in single-target displays versus accuracy for low-salience targets in dual-target displays; here, however, analyses were restricted to trials on which the high-salience target was identified first. When this modified calculation is applied to Fleck et al.’s (2010) data, significant SSM rates are apparent in both Experiment 3 (15-second time limit) and Experiment 5 (30-second time limit).

Global analyses across conditions reveal that there are overall effects of motivation for all measures of accuracy, but this main effect of motivation appears to be entirely driven by dramatic increases in accuracy in the condition that included both motivation and feedback. Paired comparisons reveal no differences between motivated and non-motivated conditions that did not include feedback and no differences between feedback and no-feedback conditions that did not include motivation. Motivation or feedback alone do not appear to alter multiple-target search performance, but the combination of these two factors produces higher accuracy rates overall and for each type of trial and target.
Why would motivation and feedback only be productive in combination? In the case of motivation without feedback, it is possible that, participants had no awareness of how they were performing. Despite being motivated and attempting to search thoroughly, multiple-target search is a difficult task, and they may not have realized they were missing targets. Observers are sensitive to the statistics of their environment, and in the context of multiple-target search, adjust their strategies in accordance with particular distributions of target frequencies (e.g., Cain et al., 2012). The group receiving both motivation and feedback may have effectively employed the feedback to adjust expectations about target frequencies while the group without feedback assumed that targets occurred less frequently than the did. Feedback alone may not have resulted in accuracy increases because participants were not concerned with how they were performing if they had no reason to be (e.g., monetary incentive). Therefore, only in combination could motivation and feedback serve to improve performance.

6.4. Acknowledgements

For helpful conversation, we thank Elise Darling, Deepu Murty, Emma Dowd, and Joe Volosky. This work was partially supported by the Army Research Office (#54528LS) and partially through a subcontract with the Institute for Homeland Security Solutions, a research consortium sponsored by the Resilient Systems Division in the Department of Homeland Security (DHS). This material is based upon work supported by the DHS under Contract No. HSHQDC-08-C-00100. Any opinions, findings, and
conclusions or recommendations expressed in this material are those of the authors and
do not necessarily reflect the official policy or position of DHS or of the U.S.
Government. The study is approved for public release.
7. Conclusion

Visual search can inform a variety of cognitive processes, and the work in this document focuses on the implications of research with regard to experience, learning, context, and motivation. Search is also essential for numerous real-world processes and careers, and thus, the study of visual search presents opportunities for both the advancement of cognitive theory and its applications. From the perspective of cognitive theory, visual search may offer a means to explore an intermediary area of processing between basic perception and broad expertise. Traditional theories of perceptual learning focus on improvements in low-level sensory discriminations, which are essential to search, but cannot account for all improvement in visual search. Visual search also relies upon higher-level, non-specific mechanisms, which can be trained directly but may also be improved through acquisition of a more global form of expertise. Furthermore, the cascade of processes involved in successful search tends to be sensitive to environmental influences that can enhance or diminish performance via motivation, anxiety, and related emotional states. Finally, because visual search is a critical applied skill for careers such as radiology and airport security, this work aims to apply insights gained from basic research to relevant occupational settings.

Historical studies and key cognitive theories of search have formed a solid framework for further exploration of exactly what guides visual search performance. Building upon the existing literature, the studies detailed in Chapters 2-6 each discuss a
manner by which visual search abilities or performance differ, or change, with respect to a specific factor. The prospects of cognitive flexibility and neuroplasticity are enticing to researchers and valuable for society at large because of a wide array of benefits that could result from cognitive enhancement.

Traditionally, research has aimed to minimize variability within participants, attempting to reduce potential confounding differences and assess the impacts of controlled manipulations. However, individual differences are readily apparent across all domains, and there has been a recent interest in exploring these differences and cognitive enhancements within unique populations. Likewise, variations associated with contextual and environmental factors can result in dramatic differences in cognition. Exploring such factors can inform interactions between streams of cognitive processing and provide suggestions for optimal structures within workplace environments. By examining variation in visual search abilities and performance associated with a wide spectrum of influential factors, the work discussed here informs the sensitivity of cognitive processes and provides insight into the complex relationship between the searcher, the searcher’s prior experience, and the environment in which she searches.

7.1. Experience and learning

The first segment of this work examines three scenarios by which visual search performance varies in accordance with learning and/or experience: 1) radiological training and practice, 2) extensive experience playing action video games, and 3) direct
training on a visual search task in the laboratory. In each scenario, there are differences in performance between the experienced and inexperienced groups/conditions, but what is the nature of these differences, and are they driven by the same mechanisms?

Each type of experience represents a unique manner through which visual search learning may occur:

1) **Radiologists** undergo extensive, focused training on visual search and are explicitly informed about optimal strategies and common sources of error. Their careers, in large part, require conducting thousands of meticulous searches of critical importance.

2) **Video game players**, specifically avid action video game players, participate in a recreational activity that has recently been associated with enhanced cognition, especially abilities related to perception and attention. They have no formal training in visual search, and their experiences playing video games are, at best, loosely aligned with the experience of conducting a focused visual search.

3) **Participants trained on a task in the laboratory** are essentially random individuals with no characteristics presumed to be associated with visual search prowess. Instead, they are repeatedly exposed to a specific task in a controlled laboratory environment and typically demonstrate improvement on the trained task.
Visual-search learning involves higher-level changes than basic-perception learning, and higher-level changes can be more complicated to untangle than low-level changes associated feature discrimination. By examining three types of visual search experience and assessing behavioral and/or neurocognitive markers of performance, the work described in the first segment of this document provides insight into mechanisms driving improvement and has implications for whether there are commonalities between sources of improvement related to experience.

7.1.1. Professional searchers and career-based experience

Extensive experience conducting visual searches in careers such as radiology likely allows for the acquisition of enhanced visual search skills, but even experienced radiologists commit errors, and the rates of false-negatives within radiographs containing multiple targets reveal a particular weakness. Radiological researchers have attempted to identify the sources of these errors, or subsequent search misses (SSMs), using radiologists as participants and real radiographs as test stimuli (see Berbaum, 2012 for a review), but the same types of errors are observed among non-professional populations searching non-medical displays as well. Laypersons, who possess neither focused experience searching nor an awareness of the SSM pitfall, demonstrate a similar decline in accuracy for second targets when searching simplified displays in laboratory settings (e.g., Fleck et al., 2010). Prior to the study discussed in Chapter 2, however, few studies allowed for the direct comparison of professional and non-professional searchers.
on the same task, and thus, it was unclear whether implications drawn from studying one population could be applied to the other.

7.1.1.1. Differences in performance between radiologists and laypersons

The work described in Chapter 2 allows for the comparison of performance between practicing radiologists and laypersons on the same multiple-target search task. A broad analysis of the data indicates that, perhaps surprisingly, radiologists and layperson searchers did not differ in overall search accuracy; the percentage of trials completed correctly (no misses, no false alarms) was not significantly different between the groups. Additionally, both radiologists and laypersons exhibited significant rates of SSMs. One substantial difference between the groups was that radiologists spent significantly longer searching than undergraduates, a common difference observed between professional and non-professional populations (e.g., Biggs, Cain, Clark, Adamo, & Mitroff, 2013; Jackson, Clark, & Mitroff, 2013). Despite spending longer assessing each trial, however, radiologists still produced SSMs on this simplified task, in line with the related errors observed within radiological searching. Recent efforts have attempted to untangle the causes of SSMs (e.g., Cain, Adamo, & Mitroff, 2013), but are radiologists committing these errors for the same reasons?

7.1.1.2. Causes of multiple-search errors

Researchers in academic radiology have theorized about the causes of multiple-target search errors but have struggled to find conclusive evidence in support of a
common underlying mechanism. Some proposed causes have included 1) satisfaction (early termination of search after the identification of one target item; e.g., Tuddenham, 1962; Smith, 1975), 2) perceptual set (after finding a target of a specific type, it is easy to miss a target of a different type; e.g., Berbaum et al., 1991), 3) resource depletion (the identification of one target consumes cognitive resources required for additional target detection; e.g., Cain & Mitroff, 2013); 4) scanning errors (the second target is never fixated; e.g., Berbaum et al., 1996); 5) recognition errors (the second target is fixated but not for long enough to be identified as a target; e.g., Samuel, Kundel, Nodine, & Toto, 1995); and 6) decision errors (the second target is fixated for long enough to be considered but is ultimate dismissed as a non-target; e.g., Nodine & Kundel, 1987).

In a recent study using laypersons as participants, Cain and colleagues (2013) used eye-tracking and manipulated specific parameters within a multiple-target search task in an effort to identify how these proposed sources of errors may systemically contribute to SSMs. The largest source of errors arose from scanning errors, in which the second target was never fixated. However, in the study described in Chapter 2, a substantial difference in the behavioral markers between radiologists and laypersons suggests that SSM errors among radiologists may arise from an alternative cause.

7.1.1.3. Target-detection order biases

In the majority of studies of dual-target search within layperson populations, dual-target trials consist of one high-saliency target and one low-saliency target.
Participants almost always (~80% of the time) identify the high-salience target first, and because SSM errors are associated with errors identifying a second target, SSM rates are typically calculated as the difference between accuracy for low-salience targets on single- vs. dual-target trials. However, a striking difference between radiologists and laypersons (Chapter 2) emerged in that radiologists did not demonstrate the common layperson bias to identify the high-salience target first. In fact, radiologist participants were equally likely to identify either the low-salience or high-salience target first.

A large portion of errors in the Cain et al. (2013) eye-tracking study stemmed from scanning errors, in which the low-salience target was never fixated, and analyses of participants’ scan paths revealed no consistent search pattern from trial to trial. The bias for detecting high-salience targets first and the lack of methodical search paths among laypersons suggests that they are likely initiating a search with a broad survey of the array, identifying targets that are easy to spot (high-salience) and then continuing to search for additional targets perhaps haphazardly. Not only did the radiologists demonstrate no bias toward detecting high-salience targets first, they also reported (via post-experiment questionnaires) that they employed identical, strategic scan paths when searching each trial. Accordingly, it follows that the radiology participants would be equally likely to identify either the high- or low-salience target first.

It is certainly possible that, had eye-tracking been employed, the SSM errors of radiologists would also be categorized as scanning errors, in that the second targets were
never fixated. Given the scan paths implied by detection order as well as the radiologists' self-reported methodical search paths, however, another possibility may be related to the trial time limits within the paradigm. Among layperson populations, participants rarely, if ever, exceed the 15-second time limit per trial (<1% of trials), and mean response times fell well below the time limit. Radiologists, on the other hand, frequently exceeded the trial time limit and had significantly longer response times, even on trials completed before time expired. In accordance with standard procedures (e.g., Fleck et al., 2010), all trials on which the time limit was exceeded were discarded from analyses. Thus, radiologists did commit SSMs on trials self-marked as completed; however, their tendency to exceed trial time limits and significantly longer response time suggests the radiologists may have not actually finished searching and were simply trying to comply with the time-limit directive.

In sum, while SSMs occur among both professional and non-professional groups, the results of Chapter 2 suggest that these populations are likely searching quite differently, and evidence regarding underlying causes of SSMs cannot necessarily be directly applied from one population to the other.

### 7.1.1.4. Additional recent work with professionals in visual search

Another career that relies on successful searching is airport security, and recent work has explored differences between Transportation Security Administration (TSA) officers and laypersons on both single- (Biggs, Cain, Clark, Darling, & Mitroff, 2013) and
multiple-target (Biggs & Mitroff, 2013) search tasks. On simplistic single-target searches, in which participants indicated whether a target item was present or absent, TSA officers exhibited significantly higher rates of accuracy across all set sizes but also spent longer searching each trial than laypersons. Further analyses revealed that a speed/accuracy tradeoff might be responsible for driving differences in performance within laypersons, as response time accounted for a large percentage of within-group variability in accuracy. Similar results were found among less-experienced TSA officers (those with less than three years of experience). Among experienced TSA officers (those with more than six years of experience), however, response time was not a significant predictor of accuracy. Instead, the consistency of time spent searching from trial to trial, was the most significant predictor of accuracy within this group (Biggs et al., 2013). These results have implications for both the role of causality in superior performance observed among expert searchers and the fact that acquired strategy plays a substantial role.

The errors associated with multiple-target search pose a concern among TSA officers as well, as airport luggage could contain more than one illegal item (a water bottle and a gun). When TSA officers were compared to laypersons on the same multiple-target search task employed in Chapter 2, there were little differences in accuracy between the groups, and both populations produced SSMs (Biggs & Mitroff, 2013). Additionally, like radiologists, TSA officers did not demonstrate the typical layperson bias of detecting the high-salience target first in dual-target trials.
Furthermore, trial-to-trial consistency of response time was a large predictor of accuracy among the TSA officers, again suggesting that experts employ superior strategies. Unfortunately, due to limited access to radiologists and a very small sample size, there was not sufficient power to perform similar assessments on the data presented in Chapter 2, but similar results between the TSA officers and radiologists is consistent with a strategic explanation for performance differences observed in expert populations.

7.1.1.5. Accounting for global differences between expert populations

A standard method of assessing performance differences associated with expertise involves comparing an expert population to laypersons; however, several inherent differences between expert and novice populations often raise concerns about potential confounds. In addition to differences in age, education, and levels of motivation, experts almost always spend significantly longer assessing trials than laypersons (e.g., Clark et al., under revision; Biggs et al., 2013). In order to circumvent concerns about expert differences arising from a simple speed/accuracy tradeoff, a recent study compared two expert populations, both expected to exhibit significantly longer response times than laypersons, but each with expertise in a separate area of visual processing (Jackson, Clark, & Mitroff, 2013).

Trained orthodontists may be better able to detect subtle facial asymmetries than general dentists or laypersons (Kokich, Kiyak, & Shapiro, 1999), as any alterations in the structure of the jaw and teeth can result in dramatic changes in facial structure as a
issue; as such, orthodontists receive advanced training in assessing facial symmetry. Jackson et al. (2013) compared orthodontists, laypersons, and TSA officers on the same facial-symmetry assessment task. TSA officers spend substantially longer searching than laypersons on visual search tasks (e.g., Biggs et al., 2013), but have no special training evaluating facial symmetry, and thus, may serve as a matched control group for speed differences. Indeed, Jackson et al. (2013) found that orthodontists were significantly more accurate in assessing facial symmetry than both laypersons and TSA officers, and while the orthodontists’ response times were significantly longer than those of laypersons, they were equivalent to the response times of TSA officers. Accordingly, the TSA officers served as an effective control group, indicating that the superior facial symmetry assessment seen among orthodontists was likely not driven by a longer time spent assessing each trial. Thus, when speed/accuracy tradeoffs present a concerning difference between expert populations and laypersons, comparing experts to other experts may allow for greater control over this factor.

7.1.2. Explaining enhanced cognition among action video game players

Spawned by Green and Bavelier’s (2003) influential study demonstrating generalized cognitive benefits of avid action video game players (VGP), there has been a surge in related research over the past decade. A recent meta-analysis (Powers, Brooks, Aldrich, Palladino, & Alfieri, 2013) assessed 118 studies of VGP’s abilities, involving 569 experimental tasks, spanning nearly every cognitive domain. The prospect of
generalized learning (from video games to a wide variety of other tasks) is particularly interesting to researchers because training on one task is rarely transferable to another. While video game research has stirred the scientific community and sensationalized the public, the validity of conclusions drawn from these studies has recently gone under scrutiny.

7.1.2.1. Concerns regarding research methodology and conclusions

According to Boot, Blakely, and Simons (2011), all studies of enhanced cognition to date (whether training or correlational) suffered from at least one substantial methodological shortcoming that may undermine the results. In line with motivational concerns inherent in comparisons between career experts and novices, the authors suggest that motivation could also play a role in driving the performance benefits seen in VGPs. If a participant is aware that he/she is partaking in an experiment because of video-game prowess, he/she may be particularly motivated to prove his/her skills. Because of this issue, the authors recommended employing covert recruitment strategies, wherein there are no requisite characteristics for participation in the study, and administering a questionnaire to determine VGP status following the cognitive task. (The study discussed in Chapter 3 had in fact employed this covert recruitment strategy and was acknowledged as one of the few published studies to do so.)

In addition to causality and related issues surrounding confounding group differences, another concern with correlational studies of VGPs is that there is often no
clear relationship between playing video games and the experimental task (Boot et al., 2011). Training studies have been called into question for this reason, as well as for potential placebo effects, and many training studies have failed to be replicated (Boot et al., 2013). Furthermore, a number of training studies did not demonstrate a test-retest improvement within control groups (e.g., Ackerman et al., 2010), an unusual outcome when repeatedly exposing participants to a cognitive task. The majority of research concerned with agents of cognitive improvement tends to find performance enhancements in both the experimental and control groups, but the effects may be significantly greater within the experimental group (e.g., Appelbaum, Schroeder, Cain, & Mitroff, 2011). In the absence of a baseline test-retest improvement, it can be difficult to draw conclusions about the nature of improvement in an experimental group.

7.1.2.2. Superior searching via superior strategy

A primary source of some researchers’ skepticism regarding VGPs’ benefits is the perhaps inconceivable notion of the broad transfer of training. Learning tends to be quite specific to the task trained (e.g., Hertzog et al., 2009), and classic perceptual learning theory focuses on improvements of low-level discriminations that are not only specific to the task, but to the exact physical characteristics of the trained stimulus (e.g., Ball et al., 2002). While it is generally accepted that higher-level cognitive processes can be trained through repeated exposure to a task (and that such improvements are non-specific to the physical attributes of stimuli; e.g., Sireteanu & Rettenbach, 1995),
cognitive training via entirely separate activities violates the mechanistic propositions of learning theories.

The study discussed in Chapter 3 finds superior visual search performance among VGPs on a change-detection task, and further analyses reveal that this enhancement is likely associated with VGPs’ tendencies to search more broadly. While broader searching could be associated with the employment of a superior strategy or better visual processing (i.e., VGPs searched more broadly because they were able to absorb larger amounts of visual information), an additional analysis lends evidence against the visual-processing explanation. On each brief, serial presentation of paired images (in which one contained a slight alteration from the other), participants were to make a localization mouseclick on a location, even if they had not detected a change; then, participants indicated whether they had actually detected the change or whether response was a “guess.” It was not uncommon for participants to have unknowingly “guessed” the correct location, and both VGPs and NVGPs did so with equivalent frequency. In line with theories of implicit change detection (e.g., Mitroff, Simons, & Franconeri, 2002), enhanced visual processing would be associated with a reduction of such incidental localizations. The lack of difference between VGPs and NVGPs on this measure of implied better “vision” falls in agreement with the proposal that VGPs’ improved performance and broader searching is associated with the implementation of a superior strategy.
7.1.2.3. Expertise and enhanced cognition

One line of research concerned with learning-related changes in cognition remains largely separate from the field of perceptual learning. Expertise is a broad concept, and the term has been used to describe superior cognitive abilities associated with universal skills (e.g., humans, in general, are face experts; Kanwisher, McDermott, & Chun, 1997) as well as skills specific to a population with extensive experience (e.g., astronauts possess extraordinary mental imagery abilities; Menchaca-Brandan, Liu, Oman, & Natapoff, 2007). The majority of research in the field of expertise is concerned with unique skills observed in a population with extensive experience within a particular area.

Cognitive differences between experts and non-experts have been established in countless domains and often related to plasticity and changes in neural architecture. When assessing images of bird and cars, the brains of respective experts show enhanced activation in the fusiform gyrus (Gauthier, Skudlarski, Gore, & Anderson, 2000), a region also associated with face recognition (Kanwisher, et al., 1997). Taxi drivers, who acquire unique spatial-navigation and landmark-recognition skills, have significantly larger volumes of gray matter in the posterior hippocampi (Maguire et al., 2000, 2006). While there are established neural markers associated with expertise, they involve high-level areas of complex processing and are quite different from low-level changes associated neuronal “tuning” to specific features.
Likewise, the cognitive mechanisms believed to underlie changes related to expertise are non-specific enhancements in processing within the appropriate domain. Expert Scrabble players have superior verbal and visuospatial abilities (Halpern & Wai, 2007); expert chess players are better at “chunking” and pattern recognition (Chase & Simon, 1973). As is the case with video game players, the distinction between self-selection and the effects of learning may be unclear, but the differences apparent in Scrabble and chess players are described as enhancements in high-level processing, and such conclusions are largely accepted by the scientific community.

While a similar conclusion about video game players may not seem as sexy as the idea that gaming alters basic perceptual processing, explaining gamer benefits in terms of enhanced application of high-level strategies may be more plausible. To date, most, if not all, studies with video-game players are wrought with methodological flaws (e.g., Boot et al., 2011), and evidence in support of “better vision” is, at best, questionable. The results detailed in Chapter 3 find support for superior strategies among video game players, and investigating cognitive enhancements within this population is a worthy endeavor. However, until research can demonstrate a reliable relationship between gaming and basic vision, any benefits associated with video-game playing are better described as expertise than “generalized perceptual learning.”
7.1.3. Laboratory-based practice on a specific visual search task

Because visual search involves both low-level sensory processing and higher-level cognitive control and decision-making processes, learning could occur at multiple stages within the stream of processing. The populations discussed above exhibit superior, or different, search performance because of global changes related to the strategic use of attentional resources. While the radiologist population in Chapter 2 did not demonstrate marked improvement on the task, radiologists do possess specialized expertise with regard to medical images (e.g., Harley et al., 2009). Because of the perceptual variability from X-ray to X-ray, and the advanced anatomical knowledge required for assessing radiographs, it is unlikely that radiological expertise stems from enhanced basic sensory abilities. Similarly, the analyses in Chapter 3 suggest that VGPs’ superior searching is related to better vision.

Broad long-term experiences may enhance strategy, but training on a specific search task could alter lower-level processes and result in improvement as well. When directly trained on a visual search task in a laboratory environment, response times become significantly faster (e.g., Sireteanu & Rettenbach, 1995), and the work in Chapter 4 used event-related potentials (ERPs) to assess how various stages of processing contribute to the observed improvements.
7.1.3.1. Electrophysiological investigations of improvement in visual search

Two recent studies have used event-related potentials (ERPs) to explore mechanisms underlying changes in search efficiency (An et al., 2012; Hamame et al., 2011). Both studies required participants to undergo direct training on a conjunction (or “inefficient”; e.g., Treisman & Gelade, 1980) search task, in which the target item shares more than one physical characteristic with distractor items, and thus, does not immediately “pop out” at the observer. For instance, if searching for a purple triangle within an array of distractors consisting of triangles and squares that are both purple and green, there is no single characteristic to differentiate the target from distractors. In such cases, several fixations are typically required before observers can successfully locate the target item.

Conjunction searches are typically characterized by increases in response time as set size increases (e.g., Treisman & Sato, 1990), which is associated with the requisite sequential, or serial, searching of individual items in order to locate the target (e.g., Snodgrass & Townsend, 1980). Feature (or “efficient”) searches, on the other hand, allow for simultaneous, or parallel, searching and do not result in longer response times with larger set sizes (e.g., Koch & Ullman, 1985); because the target possesses a unique feature to distinguish itself from distractors, it immediately captures attention, regardless of the number of additional items in the array. While search slopes for conjunction tasks can begin to resemble characteristically parallel search slopes, physiological evidence
suggests they remain more demanding of attentional processing than features searches (Leonards, Rettenbach, & Sireteanu, 1997).

After training on conjunction search tasks, both An et al. (2012) and Hamame et al. (2011) found increases in the amplitude of the N2pc (an ERP component associated with attentional orienting; e.g., Luck & Hillyard, 1994a) but no changes in the latency of the N2pc. These results suggest that practice can enhance processing related to the attentional allocation of the target but does not facilitate faster attentional orienting. Thus, response time decreases observed after practice on these conjunction searches must have been driven by alternative elements within series of processes involved in visual search. However, directing attention to targets in conjunction searches produces significant trial-to-trial variability and is more complex than in feature-singleton searches, in which attentional allocation occurs easily and rapidly. In the study described in Chapter 4, participants practiced a feature-singleton search, in which they determined whether a target ellipse was oriented vertically or horizontally. To determine which mechanisms allow for enhanced performance, electrophysiological markers of processing were compared between the first and last sessions of a five-session practice protocol.

7.1.3.2. Neural mechanisms underlying enhanced visual search

As anticipated, response time decreased significantly with practice (by ~80 ms), and accuracy remained unchanged, suggesting that the improved response times were
not the result of a speed accuracy trade off. Detailed analyses of multiple ERP markers of processing revealed that changes over the course of virtually the entire visual search process, from perceptual analysis to response preparation, collectively contributed to the observed improvement. After practice, the N1, associated with early visual processing (e.g., Mangun & Hillyard, 1991), increased in amplitude, suggesting enhanced sensory analysis. Hamame et al. (2011) did not find a change in N1 amplitudes after practice on the conjunction search task, so the early cortical changes seen in Chapter 4 may be more necessary and/or beneficial for feature search tasks than for more demanding search tasks.

The earliest temporal change in processing was observed in the N2pc. Like in An et al. (2012) and Hamame et al. (2011), there was a significant increase in N2pc amplitude, but there was also a significant latency shift, which was not observed in either of the prior studies. Temporal changes in attentional allocation may occur with feature searches but not conjunction searches either because there are different mechanisms responsible for improvement in each task, or they may be more readily observed within the feature search task because of reduced trial-to-trial variability. Either way, the analyses in Chapter 4 revealed that, after practice, the N2pc component peaked ~18 ms earlier, lending a significant contribution to the ultimate improvement in response time.
An additional change in latency was apparent in the onset of the LRP, or lateralized readiness potential, which reflects the initiation of a motor response (Coles, 1988). After practice, the onset of the LRP occurred ~39 ms faster relative to the onset of the stimulus, but approximately half of this change is accounted for by the shift in the N2pc; thus, speeded preparation for the motor response contributes approximately ~21 ms to the overall improvement in response time.

This temporal change in motor initiation may be related to facilitation of target-discrimination processes occurring after the attentional allocation phase; that is, once attention has been directed to the target, determining its orientation (i.e., vertical or horizontal) becomes easier with practice. Another index of processing, the SPCN, or sustained posterior contralateral negativity, is associated with maintenance and manipulation of information in visual working memory (e.g., Jolicoeur et al., 2006; Vogel & Machizawa, 2004), and larger SPCN amplitudes reflect greater demands on cognitive processing (e.g., Eimer & Kiss, 2010). After practice, SPCN amplitudes decreased significantly, suggesting that target analysis became easier, which likely contributed to the speeding of motor preparation for response.

Together, the latency shifts observed in the N2pc and the LRP account for ~39 ms of the ~80 ms change in response time, indicating that additional speeding of processing occurred between the initiation and execution of the motor response. This change is likely associated with increased automaticity in motor behavior with practice and
suggests that some of the improvement in response time on visual search tasks is driven by participants becoming faster to press a button. The fact that a large portion of decreases in response times can be explained by increased motor automaticity may also explain why response times on conjunction tasks change so dramatically without a reduction in the attentional demands (e.g., Leonards, Rettenbach, & Sireteanu, 1997).

The methodology in Chapter 4 allowed for a systematic analysis of mechanisms contributing to enhanced performance via direct practice on a feature-singleton search task and found significant changes in electrophysiological markers of sensory processing, attentional allocation, target discrimination, and motor preparation. Certain changes associated with sensory processing and attentional allocation were not observed in electrophysiological studies of practice on conjunction search tasks, which suggests that improvement in visual search can result from different mechanisms, depending on the task and/or the type of training. Accordingly, while the mechanisms related to improvement through laboratory-based practice on this simplistic task offer unique insight into the flexibility of cognitive processes, it is unlikely that the same mechanisms explain visual search “expertise” acquired through long-term experiences.

7.2. Context and motivation

The second segment of this work is concerned with external factors that influence visual search performance, including the framework within which a search occurs and performance-based incentives. Unlike the experiences discussed in the first section,
which may result in long-term changes in visual search abilities, situational factors can alter performance on a moment-by-moment basis, and the effects are not usually retained when the factor is removed.

Despite these differences, contextual influences on visual search are also inherently intertwined with experience. Searches conducted in occupational settings are likely taking place under motivated conditions; radiologists and baggage screeners should be motivated to search accurately not only to prevent fatalities but also for the sake of their own professional success. An overarching concern when assessing effects of experience, especially when comparing different populations, is the potential for differing levels of motivation between the groups. Such concerns may be largely alleviated by the implementation of a control task or an additional comparison to a more closely matched control group (as described in Section 7.1.1.3), but it is useful to explore exactly how motivation can impact visual search processes and performance.

Motivation has been shown to impact search performance in terms of speeded attentional deployment in basic search task (e.g., Anderson et al., 2011) and in complex scenarios such as rare-target search (e.g., Navalpakkam et al., 2009). However, motivation is not a unitary concept, and the nature of motivational incentives can induce excitement associated with reward and/or anxiety about potential failure. Motivation to avoid a punishment can be particularly stressful and promote anxiety (e.g., Davis &
While career searchers may be motivated, does such motivation have a necessarily positive influence on performance? A certain anxiety must accompany the notion that search errors could result in fatalities, and anxiety is associated with a decline in accuracy on numerous cognitive tasks, including multiple-target searches (e.g., Cain et al., 2010), which are often encountered by radiologists and airport security officers alike. The studies discussed in Chapters 5 and 6 examine the impacts of motivational reward structures on multiple-target search performance in terms of the structure of a search scenario and in combination with factors including feedback and time pressure.

7.2.1. Task framing

The conceptual framework of a situation can dramatically alter behavior, and occupational searches are conducted within unique frameworks that are not consistent across fields. Radiologists are assigned a workload consisting of X-rays to assess; they examine each case for abnormalities and are aware of the number of cases remaining in the current workload. Airport security officers, on the other hand, are assigned to work at the checkpoint for a certain time period; they examine luggage for contraband and are aware that their current search duties will be complete at the end of the shift, regardless of the number of suitcases searched. Both radiologists and airport security officers
should be attempting the same process – carefully examining each display for potentially harmful targets, regardless of the number of cases yet to be scanned or the amount of time left before the end of a shift – but does the knowledge of the remaining workload, in terms of either the number of searches or the amount of time, influence searching?

7.2.1.1. Differences in SSMs between fixed objective and fixed duration conditions

To examine the potential impacts of task structures on search performance, the study discussed in Chapter 5 compared performance between two groups of participants searching within frameworks akin to those employed in either radiology or airport security. Both groups were provided a monetary incentive for accurate searching and performance-related feedback throughout the task; however, they were aiming for different goals: one group was attempting to perform as many accurate searches as possible in a set time period (fixed duration; similar to airport security), while the other was attempting to achieve a set number of accurate searches in the shortest time period (fixed objective; similar to radiology). The two frameworks call for identical strategies and impose the same amount of “time pressure” on observers, and all measures of performance between the groups were identical as well – except SSM rates.

The only performance difference between the groups was the likelihood with which participants found the additional targets on dual-target trials, with superior accuracy for multiple targets in the fixed duration condition. This group of participants,
who were instructed to accomplish as much as possible in a specified time period found second targets more deftly than those who were instructed to achieve a specified goal in the shortest number of minutes. Unlike those in the fixed duration condition, participants in the fixed objective condition produced SSMs, showing a substantial decline in accuracy for second targets.

7.2.1.2. Interpretation of task framework effects

Because the two conditions are objectively identical in terms of optimal strategy, these results are perhaps somewhat puzzling. However, humans are, by nature, irrational beings. Humans have demonstrated a consistent propensity for irrational decision-making (e.g., Tversky & Kahneman, 1981), and decision-making is often a core component of visual search. Observers need to decide whether an ambiguous target is a target or distractor and ultimately determine when to stop searching. Such decisions become more complex when targets are more similar to distractors and when there may be more than one target within a display.

The structure of a task’s goals can produce emotional impacts, and these two frameworks may elicit different subjective experiences of stress. Because stressful motivation may inhibit cognitive performance (e.g., Lang & Bradley, 2009), it is possible that the fixed objective framework examined here induced perceptions of time pressure and/or anxiety, which have both been shown to increase SSMs (Fleck et al., 2010; Cain et al., 2011). Importantly, both conditions imposed the same amount of time pressure, as
speed is equally critical for success in either case, but the pressure may have been perceived differently between the conditions. The framework of completing a workload as quickly as possible might feel significantly more stressful than accomplishing as much as possible in a pre-determined amount of time, in which case, the results parallel those observed under experimentally induced anticipatory anxiety (Cain et al., 2011).

7.2.1.3. Feasibility of implementing a fixed duration structure in occupational settings

While accuracy is critical in occupational searches, efficiency is also required. Radiologists must assess cases in a timely manner, or life-threatening illnesses could be diagnosed too late. Airport security screeners are often faced with long lines and frustrated passengers and must facilitate reasonably efficient air travel. The results in Chapter 5 suggest that the fixed duration structure employed in airport security may promote accuracy and reduce SSMs, but could it also be implemented in radiology without inducing a decline in productivity? Medical personnel could be concerned that assigning radiologists to search in shifts, with no requirement for number of cases completed, would result in fewer X-rays assessed and longer waits for diagnoses.

However, the results of Chapter 5 find no differences between total trial time between the fixed duration and fixed objective conditions. Both task structures allowed participants to search at their own pace, but it appears that participants in both groups found a similar search speed to allow for the maximum rate of completing trials correctly. Because the fixed duration group achieved greater search accuracy without
spending a longer time searching, implementing a similar structure for radiologists may not slow medical procedures. In order to ensure the requisite volume of X-rays were assessed, radiological groups could maintain additional radiologists “on call,” in the event that the cases were not completed during the assigned shifts.

7.2.2. Dissociating effects of motivation, feedback, and time pressure

The task structures employed in Chapter 5 both contained components of motivation, feedback, and time pressure, each of which may contribute unique effects to search performance and/or interact with the other factors, and the series of studies in Chapter 6 aims to dissociate these effects. The type of motivation employed in Chapter 5, while offering a reward for superior performance, also included scoring penalties deemed necessary for the task structure, but because penalties can result in performance declines (e.g., Murty et al., 2011), no penalties were incurred by participants in Chapter 6 in order to isolate effects of rewarding motivation. Furthermore, the feedback in Chapter 5 involved a running tally of score and time elapsed, which were also necessary components of the task structure, but the feedback employed in Chapter 6 only provided information regarding correct/incorrect responses. Finally, “time pressure” in Chapter 5 was related to the amount of time allowed for completing the full experiment, but the experiments in Chapter 6 induce time pressure through time limits for individual trials.
7.2.2.1. SSM errors in all conditions

Recent studies of influences on multiple-target search performance have primarily focused on the presence or absence of SSM errors as a result of experimental manipulations (e.g., Clark et al., 2014; Cain et al., 2011). Most basic measures of accuracy (e.g., accuracy on single-target trials) are similar across conditions, and differences emerge only in relative accuracy for second targets on dual-target trials. The typical lack of variability between conditions on basic measures of accuracy could be attributed to ceiling effects (especially in the case of single-target trials containing high-saliency targets in which accuracy is usually >95%) or standard performance baselines that are simply not impacted by the nuanced manipulations. These studies proposed that the effects of experimental manipulations such as task framework and anticipatory anxiety result only in changes in SSM rates because of the unique cognitive challenges associated with searching for multiple targets. Accordingly, the experiments described in Chapter 6 were intended to allow comparison of which combinations of factors (i.e., motivation, feedback, and time pressure) did and did not produce SSM errors. Two conditions were near replications of prior experiments (e.g., Fleck et al., 2010): one produced SSM errors (time pressure, 15-second time limit; Experiment 3) and the other did not (minimal time pressure, 30-second time limit; Experiment 5). Similar results were expected in Chapter 6, and the additional manipulations (i.e., motivation and
feedback) were implemented to allow comparison to the anticipated replication of the effects of time pressure.

Contrary to predictions, SSM errors rates were at least marginally significant across all eight conditions and highly significant in most. Additionally, there were dramatic differences across the basic measures of accuracy that typically remain unchanged with contextual manipulations, and thus, the analyses in Chapter 6 examine changes in all measures of accuracy rather than focusing on SSM error rates alone.

7.2.2.2. No effect of time pressure on multiple-target search accuracy

The results in Chapter 6 found no impact related to the time pressure manipulation, wherein participants were limited to a limit of 15 seconds per trial. This lack of variation in multiple-target search performance with respect to time pressure is surprising in light of prior work, which suggests that time pressure may be a critical component involved in inducing SSM errors. Fleck et al. (2010) examined multiple-target search performance among laypersons with regard to several manipulations and found that SSM errors emerged when trials had a time limit of 15 seconds (Experiment 3) but not when trials had a time limit of 30 seconds (Experiment 5). The non-motivated conditions without feedback in Chapter 6 were identical to Fleck et al.’s (2010) experiments (except than the no-time-limit condition had no time limit at all rather than an extended time limit of 30 seconds). However, the differences between the results can be explained by a minor difference in the method of calculating SSM rates.
Fleck et al. (2010) was the first study to employ the current methodology in order to examine multiple-target search errors in non-professional populations, and since this time, additional investigations have garnered further insight and resulted in some modifications in methods and measures. In Fleck et al. (2010), SSM rates were calculated as the difference between accuracy for low-salience targets in single-target displays versus accuracy for low-salience targets in dual-target displays. SSMs are, by definition, errors in identifying the second target in a display, and in these paradigms (in which dual-target trials contain one high-salience target and one low-salience target), low-salience targets are almost always (~80% of the time) the second target identified in dual-target trials. However, Fleck et al.’s (2010) SSM calculations included accuracy for low-salience targets on all dual-target trials, including the ~20% of dual-target trials on which the low-salience target was actually the first target identified. More recent explorations of SSM errors within related paradigms (e.g., Cain & Mitroff, 2013) have tweaked this calculation, limiting the comparison to only the trials on which the high-salience target was identified first, in order to more accurately capture errors associated with second targets.

When this modified calculation is applied to Fleck et al.’s (2010) data, significant SSM rates are apparent in both Experiment 3 (15-second time limit) and Experiment 5 (30-second time limit), contrary to the conclusions discussed in the study. Furthermore, there are no significant differences on any measure of accuracy or response time.
between those data and the associated data in Chapter 6. The *time pressure* condition in Chapter 6 replicates Fleck et al.’s (2010) Experiment 3, and the results in the *no time pressure* condition are equivalent to Fleck et al.’s (2010) Experiment 5 (i.e., there are no differences between a 30-second trial time limit and no time limit). The effect of *time pressure* described by Fleck et al. (2010) was a significant factor responsible for motivating the current experiments, but it appears that this original conclusion may have been misleading, as time pressure produces no effect on accuracy for genuine “second” targets.

7.2.2.3. The combination of motivation and feedback promotes multiple-target search accuracy

Global analyses across conditions revealed that there are overall effects of motivation for all measures of accuracy, but this main effect of motivation appears to be entirely driven by dramatic increases in accuracy in the condition that included both motivation *and* feedback. Paired comparisons reveal no differences between motivated and non-motivated conditions that did not include feedback and no differences between feedback and no-feedback conditions that did not include motivation. Motivation or feedback alone do not appear to alter multiple-target search performance, but the combination of these two factors produces higher accuracy rates overall and for each type of trial and target.

Why would motivation and feedback only be productive in combination? In the case of motivation without feedback, it is possible that, participants had no awareness of
how they were performing. Despite being motivated and attempting to search thoroughly, multiple-target search is a difficult task, and they may not have realized they were missing targets. Observers are sensitive to the statistics of their environment, and in the context of multiple-target search, adjust their strategies in accordance with particular distributions of target frequencies (e.g., Cain et al., 2012). The group receiving both motivation and feedback may have effectively employed the feedback to adjust expectations about target frequencies while the group without feedback assumed that targets occurred less frequently than the did. Feedback alone may not have resulted in accuracy increases because participants were not concerned with how they were performing if they had no reason to be (e.g., monetary incentive). Therefore, only in combination could motivation and feedback serve to improve performance.

7.2.2.4. Implications for occupational protocols

If search performance is enhanced with the combination of motivation and feedback, can workplace protocols benefit from employing these factors? Career searchers are, presumably, already motivated, and advancement structures within both radiology and airport security allow for performance-based promotions. Feedback, however, is not typically provided, as there is ground truth to which to compare searchers’ assessments. Radiologists may be informed of inaccurate assessments when a later medical procedure reveals a previous error for a particular patient, but such
feedback occurs long after the mistake, making it difficult to remember, and modify, the
cognitive processes responsible for the error.

Airport security screeners do receive some feedback for their performance on
*threat image projections*, or TIPs – false target items that are randomly inserted into
passengers’ bags. When a TIP is present and the officers digitally indicate the suspicious
bag, the system reports that they have correctly identified a TIP item, and similar
feedback is provided for missed TIPs. The explicit goal of TIPs, however, is not to
provide feedback but to counter the effects of low target prevalence and maintain
alertness. There are also potential issues surrounding the use of TIPs; some officers
report that TIPs are easily distinguishable from real items in that they appear in
unrealistic locations and orientations. Thus, searching for TIPs may be cognitively
separable from searching for true target items, in which case the feedback provided
would be irrelevant. Unfortunately, implementing feedback in occupational searches
may be virtually impossible. One possibility would involve more than one individual
searching each X-ray and comparing any differing assessments, but such an endeavor
would require essentially doubling the workforce, a cost that would likely exceed the
benefits.

7.3. *Closing thoughts*

Collectively, the content discussed here informs core questions about cognitive
flexibility: How and why might cognitive processes change? What types of experiences
enhance cognition, and are there multiple means by which to do so? Are there
differences between changes associated with simple versus complex processes? What
other factors can modulate cognitive performance, and how are these factors associated
with experience and learning? How can we apply the knowledge gained from basic
research to improve cognition in the real world?

7.3.1. The nature of learning in visual search

Visual search is the product of low- and high-level cognitive processes, both of
which can be altered through experience and training. Enhancement in any one of the
mechanisms involved in searching will produce an overall improvement in
performance. Thus, superior searching observed after different types of experiences and
training cannot necessarily be attributed to a common cause. Occupational searchers and
video game players have extensive experience with perceptual activities that provide
exposure to visual stimuli with infinitely variable physical characteristics. These long-
term perceptual experiences cannot train sensory processing, but instead, serve to
enhance strategic allocation of attention, which facilitates visual search abilities. Low-
level visual skills can be trained through repeated exposure to a specific stimulus, but
laboratory-based training on visual search tasks reveals alternative routes for
improvement. Practicing a visual search task results not only in changes in basic sensory
processing but also in attentional allocation, target discrimination, and motor-response
initiation. The means by which visual search can improve are as numerous as the processes it entails.

7.3.2. Influences on multiple-target search errors

Errors associated with searching for multiple targets occur among both professional searchers and laypersons but are driven by different causes. Laypersons tend to search haphazardly, failing to employ a consistent method or strategy. They excel at identifying salient item but struggle to locate less obvious targets. Despite efforts to search for additional targets, cognitive resources are expended during the identification of one target and limit attention during the remainder of the search process. Radiologists, on the other hand, are methodical searches who are equally adept at identifying high- and low-salience targets. They spend significantly longer searching and may commit errors because of time constraints; however, the causes of multiple-target search errors among radiologists remain largely unclear.

A variety of factors alter multiple-target search accuracy among laypersons. Minor manipulations, such as the framework of a task, that do not modify accuracy for single-target searches can induce unique effects specific to second-target accuracy. Stronger manipulations can improve accuracy for single- and multiple-target searches alike, and the combination of motivation and feedback produces especially high levels of accuracy.
References


Eimer, M., & Kiss, M. (2010). The top-down control of visual selection and how it is linked to the N2pc component. *Acta Psychologica, 135*, 100–102.


Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*, 368–373.


Biography

Kait Clark was born in Philadelphia, Pennsylvania, on January 2, 1986, the first and only child of parents Margaret (Peggy) Ann Therwanger Clark and Dennis Patrick Clark. She grew up in Philadelphia and attended Nazareth Academy Grade School for her early education and Mount Saint Joseph Academy in nearby Flourtown, Pennsylvania, for high school, graduating in June 2004. Kait then enrolled at Saint Joseph’s University, where she majored in Psychology and conducted research in the laboratories of Drs. Patrick Garrigan and John Jewell. She graduated Summa Cum Laude in May 2008, and shortly thereafter, moved to Durham, North Carolina, to begin her graduate studies at Duke University, where she worked under the mentorships of Drs. Stephen Mitroff and Marty Woldorff.

Publications

Referred Journal Articles


Clark, K., & Mitroff, S. R. (in preparation). I knew you were going to miss that: Predicting future visual search performance from initial search abilities.


Clark, K., Zucker, N. L., Watson, K. K., & Mitroff, S. R. (in preparation). Revealing individual differences in visual search: ADHD leads to seeing too little and anorexia nervosa leads to seeing too much.

**Book Chapters & Peer-Reviewed Conference Papers**


Honors & Awards

Vision Sciences Society Travel Award, 2012
Object Attention, Perception, and Memory Travel Award, 2011
Fellowship to Summer Institute in Cognitive Neuroscience, UC Santa Barbara, 2009
James B. Duke Fellowship, Duke University, 2008 – 2013
Departmental Honors, Saint Joseph’s University, 2007 – 2008
Honors Program, Saint Joseph’s University, 2004 – 2008
Dean’s List, Saint Joseph’s University, 2004 – 2008
Presidential Scholarship, Saint Joseph’s University, 2004 – 2008
Kelly Scholarship, Saint Joseph’s University, 2004 – 2008

Professional Affiliations & Memberships

Vision Sciences Society, Member, 2007 –
Phi Beta Kappa, Member, Saint Joseph’s University Chapter, 2008 –
Psi Chi, Vice President, Saint Joseph’s University Chapter, 2007 – 2008
National Society of Collegiate Scholars, Member, 2005 –