UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL AND HUMAN SCIENCES

Psychology

Search Behaviour in Real World Tasks: The Influence of Working with a Partner

by

Charlotte Riggs

Thesis for the degree of Doctor of Philosophy

October 2016

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL AND HUMAN SCIENCES

Psychology

Thesis for the degree of Doctor of Philosophy

Search Behaviour in Real World Tasks: The Influence of Working with a Partner

Charlotte Riggs

Visual search in the real world is a task that we all engage in on a regular basis, for example, looking for a set of keys in a living room or a mobile phone in a handbag. Visual search is typically examined in a laboratory setting, where individuals search for simple targets on a computer display. However, many real-world search tasks are more complex and challenging, and may require searchers to work together. This thesis explored some of the factors influencing the effectiveness of search behaviour in complex real world tasks. Of key interest was how searching may be affected by searching alone or with a partner. Chapter one provides a general introduction and review of previous research. A number of challenges to searching in the real world are revealed. In regards to searching with a partner, in the context of airport X-ray security screening on a computer display, Chapter 2 reveals that searchers should be encouraged to actively ignore another searcher's target and quality of search. In the context of searching for clues to a crime or Improvised Explosive Devices (IEDs) in patrol on a large area of grassland, Chapter 3 reveals that pairs should be encouraged to work independently of each other to avoid any costs in coordination. In the context of rummage searching both visually and physically for threatening items in a large-scale house environment, Chapter 4 reveals that previous findings may be true in novice searchers, but that the key to coordination in pairs is training searchers to double check each other. Chapters 3 and 4 also form and validate two novel analytic frameworks, which can be utilized for exploring search in real world scenarios. Chapter 5 considers the implications to real world search tasks and suggests future research.

Contents

List of Tables

List of Figures

Declaration of Authorship

I, *Charlotte Riggs* declare that the thesis entitled *Search Behaviour in Real World Tasks: The Influence of Working with a Partner* and the work presented in it are my own and has been generated by me as the result of my own original research. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University;
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have consulted the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- Either none of this work has been published before submission, or parts of this work have been published as:

Donnelly, N., Riggs, C. A., Mann, C. M., Godwin, H. J., Menneer, T., & Liversedge, S. P. (2013). Specialist search human factors and training - improving rummage search performance: the role of threat briefings. Technical Report for Dstl.

Donnelly, N., Riggs, C. A., Mann, C. M., Godwin, H. J., Menneer, T., & Liversedge, S. P. (2013). Specialist search human factors and training - improving rummage search performance: the influence of experience and training. Technical Report for Dstl.

Godwin, H. J., Menneer, T., Riggs, C. A., Taunton, D., Cave, K. R., & Donnelly, N. (2016). Understanding the contribution of target repetition and target expectation to the emergence of the prevalence effect in visual search. *Psychonomic bulletin & review, 23(*3), 809-816.

Riggs, C. A., Cornes, K., Godwin, H. J., Liversedge, S. P., Guest, R & Donnelly, N (in press). The importance of search strategy for finding targets in open terrain. *Cognitive Research: Principles and Implications.*

Signed: …………………………………………………………………………..

Date:……………………………………………………………………………..

Acknowledgements

First and foremost, I would like to thank my primary supervisor, Professor Nick Donnelly, for his advice, support, and the great deal of time he spent providing feedback. Without his help, this thesis would not have been possible. Furthermore, I would like to thank him for guiding me and inspiring my interest in Psychology for many years prior to my PhD, and also for providing me with the opportunity to carry out this project. I would also like to thank my second supervisor, Professor Simon Liversedge, for the many pep talks over the years, but also for vital advice and feedback on chapters and writing for publication. Furthermore, my thanks must go to Dr Hayward Godwin, for sharing his knowledge, ideas and skills, which both shaped my approach to research and helped me to prepare for this task. I have also benefitted from the help and advice of Dr Natalie Mestry, who has always been willing to answer any questions I might have. My thanks also go to collaborators that I have worked with during my PhD, including Dr Tamaryn Menneer, Dr Katherine Cornes, Dr Richard Guest, Dr Sarah Smith, Mike Boardman, Carl Mann and Liam Riddy.

I am grateful to everyone that I have shared room 3109 with, for being both my colleagues and my friends. Thank you for providing the necessary humour to get me through the tougher times. A special thanks goes to Gemma and to Lewis, for always lending an ear and providing an output for frustration, and to Fede, for being my friendly office neighbour, offering daily advice when needed.

I am also grateful to friends outside the University, for their support and encouragement, while providing much needed escapism and helping to take my mind off of work. A special thanks goes to Katy, for not only proof reading chapters near the end, but for believing in me from the very start, and offering support and advice when I needed it the most. I am also grateful to my housemate Kayleigh, for always being there. A special thanks also goes to Pete, for cheering me on whilst showing patience and understanding during my write up period. Finally, my sincere thanks go to my family, for their continued faith in me and for their endless support, both emotionally and financially.

The research in this thesis was funded by the University of Southampton's Vice Chancellor Scholarship Scheme.

Abbreviations

ESW	effective search width
IED	improvised explosive device
min	minute
ms	milliseconds
ROI	region of interest
RS	rummage search
s	second
SD	standard deviation
SEM	standard error of the mean

Chapter 1

Introduction

How individuals perform visual search has been a topic of significant research interest over recent decades. By visual search I refer to the processes of selection and guidance that allow for targets to be located in the visual field and identified. Visual search is typically explored by running experiments on computer screens. The conclusions from these studies are often extrapolated to provide explanations for searching beyond the computer screen. Searching in the real world *can* be a simple task, but more often it is a complex and challenging one. To address the problem of how searchers might behave in a number of real world search tasks, as explored in this thesis, what is already known about search (in terms of what conditions lead to accurate or poor search) and how search is influenced by the demands of searching in the real world needs to be considered.

Models of visual search consider accuracy and speed of detection in basic visual stimuli. Exploring these models reveals what happens when more or less distractors are present, and how features such as colour, shape and orientation make search more or less efficient. Although these models use simple stimuli, and can be seen as naïve in relation to the complexity of real world search tasks, they provide a fundamental platform for how basic search is guided. The quality of this guidance being affected by a range of factors from the presence of contextual information, clutter and the number of targets sought.

As the task demands associated with searching become more complex, search is influenced by biases and decisions that influence search strategy and the goals of search itself. In this case, search can be as much about ensuring target absence as much as determining target presence. To explore real world search tasks and provide an account of searching in the real world, the above factors need to be considered. These are the issues that will be explored in Chapter 1.

Visual Search

Visual search is a task that we engage in when looking for a set of keys in a living room, a mobile phone in a handbag, or an icon on a computer display. Visual search requires that we focus attention on a subset of all visual information. An understanding of how visual information is processed to enable search provides a background to how the visual system defines and detects targets.

Basic Experimental Data

In a typical visual search study participants are presented with a target, set amongst a number of distractors. The number of distractor items presented is referred to as the display's set size. The set size is varied over trials and both the accuracy and speed of target detection is measured and recorded (some later studies have also incorporated the measure of eye movements). Search efficiency is measured in terms of how the speed and accuracy of responses vary with set size (Duncan & Humphreys, 1989; Treisman, Sykes, & Gelade, 1977; Treisman & Gelade, 1980).

Models of Visual Search

Feature integration theory

The most well known early model of visual search is the Feature Integration Theory (FIT; Treisman & Gelade, 1980). FIT describes two separate forms of search: namely, parallel and serial search. Treisman and Gelade (1980) hypothesised that visual search could be conducted in parallel (i.e. more than one item processed at a time) on features which could be detected pre-attentively (e.g. colour, shape, orientation, movement, etc.) and in serial (i.e. one item processed at a time) on conjunctions of features (e.g. coloured shapes). When participants are asked to search for targets that can be differentiated from distractors on the basis of simple features (e.g. searching for a pink target amongst brown and purple distractors, or searching for the letter 'O' amongst 'N' and 'T' distractors; Treisman et al., 1977) results showed that the time required to perform a feature search did not increase with the number of distractors, i.e. search was efficient as the target 'popped out' of the display. Flat search functions were taken as a marker of pre-attentive search.

If targets can only be differentiated from distractors on the basis of a conjunction of features (e.g., a vertical red line embedded amongst a background of horizontal red and vertical green lines; Treisman & Gelade, 1980), then speed and accuracy decreases linearly with set size. In terms of FIT, these empirical findings were thought to indicate a serial search where target detection required the movement of attention (acting as a spotlight) to each object location.

FIT provided a framework for understanding early visual processing. Although the theory made a significant impact on research, many of its assumptions have since been challenged, as further studies revealed large variations in search efficiency that could not be explained by FIT (Duncan & Humphreys, 1989), and as evidence was found to show that efficient search could in fact occur on conjunctions of features (Wolfe, Cave, & Franzel, 1989), in direct contradiction to FIT. FIT has been updated and developed over the years (Treisman & Gormican, 1988; Treisman, 1988; 1991; Treisman & Sato, 1990) however, the notion of the spotlight of attention examining a display has remained as a crucial feature of the FIT model (Findlay & Gilchrist, 2003).

Resemblance Theory

Duncan and Humphreys (1989) suggested that the magnitude of search slopes reflects the dissimilarity between targets and the heterogeneity of distractors. They proposed that the visual field is organised and segmented in parallel into structural units that possess a shared property and that access into visual short-term memory is limited by the need for cognitive resources. When a display contains various structural units (perhaps including targets and groups of distractors), these units compete for access into the visual short-term memory. The closer the match between a target template stored in memory and a structural unit, the more resources that unit will receive. Therefore, search speed is a function of competition for access to the visual short-term memory. FIT predicts that if a target can only be differentiated from distractors on the basis of a conjunction of features, search efficiency would decrease with the number of objects present. Resemblance theory, however, predicts that set size will not influence search efficiency, but that search will be more efficient when target-distractor similarity decreases, and when distractor-distractor similarity increases (Duncan & Humphreys, 1989).

Guided Search Theory

The guided search model (GS) was originally described by Wolfe et al., (1989), but has since been modified in three separate developments (Wolfe & Gancarz, 1996; Wolfe, 1994; 2007). Wolfe et al. (1989) proposed that a search display forms an activation map, whereby features in common with the target cause activation, directing attention towards them. Search guidance is led by a master map, which receives input from feature maps containing both top-down (target-matching features) and bottom-up (e.g. contrast or saliency of the stimuli) components (Chan & Hayward, 2013).

Wolfe et al. (1989) demonstrated that when more than one defining feature is present, and search requires some form of serial inspection (i.e. when the target does not 'pop out') search for triple conjunctions (e.g. targets defined as specific combinations of colour, form and orientation) is quicker than double conjunctions (e.g. targets defined as specific combinations of colour and form). This is because the addition of a third feature provides additional information to guide attention to items that are more likely to be the target. This critical finding demonstrated that conditions could be created in which conjunctions of features could be used to allow attention to be focused only on a subset of potential targets. This meant that grouping across features must occur in parallel to identify possible targets, thereby contradicting a major assumption of FIT.

The current version of the GS model is Guided Search 4.0 (GS4; Wolfe, 2007), which has sought to incorporate top-down influences on search. The searcher's knowledge guides search to relevant areas in the scene, and a 'priority map' is created to represent the most likely locations of targets. The map is then used to direct attention to items in their order of priority, which are then selected and matched against the searcher's internal representation (Wolfe, Evans, Drew, Aizenman, & Josephs, 2016). In GS4, objects enter the search system in serial, however more than one object can be present within the search system simultaneously. In other words, GS4 suggests that search is limited by an information-processing bottleneck, which determines serial entry into the search system. Passage through the bottleneck is determined by the resemblance (similarity) of items to targets. Once through the bottleneck, the object recognition system examines the individual objects for enough evidence to determine target identity. This decision is contingent on the threshold for a target or distractor item. Therefore, although multiple objects can be simultaneously processed in parallel within the system, they can exit, or finish their journey, at different time points.

Wolfe (2014) noted, that there is a set of additional models that share aspects of the feature integration theory (Treisman & Gelade, 1980) and the guided search model (Wolfe, 2007) to varying degrees (e.g. Cave, 1999; Cohen & Ruppin, 1999; Hübner, 2001; Lee, Buxton, & Feng, 2005; Mozer & Baldwin, 2007; Tsotsos, 2011; Vidyasagar, 1999). However, although even the author (Wolfe, 2007) has raised a number of issues with the present model of GS4, its current state represents some of the most recent developments in the accounts of visual search debate and how visual search has been understood.

In addition to early behavioural measures of response speed and accuracy, advancements in eye tracking technology have meant that the recording of eye movements has become popular as a method for studying cognitive processes (Liversedge & Findlay, 2000), including visual search[1]. The question is, do these data differentiate the theoretical perspectives?

Zelinsky and Sheinberg (1997) found support in eye movements for Treisman and Gelade's (1980) FIT, revealing that fixations were longer, saccades were shorter and more eye movements were made in a serial search task than in a parallel search task; the number of eye movements made increased linearly with increases in set size as the task became more difficult and less efficient. This demonstrated a link between

[1] There are two key components to eye movements: saccades and fixations. Saccades are the eye movements themselves, and fixations are the period of time where the eyes remain relatively still. It is during a fixation that new information is extracted from the visual display, while vision is supressed during a saccade. Saccades occur during visual search, however, the anatomy of the retina only allows detailed visual processing at the central part of the eye, the fovea. This means that when targets are complex (or small), identification requires ensuring targets are positioned on the fovea (Carrasco, Evert, Chang, & Katz, 1995; Carrasco, McLean, Katz, & Frieder, 1998; Carrasco & Frieder, 1997; Rayner, 2009). The fovea covers a region with a visual angle of two degrees, and expanding outwards from the central fovea, visual accuracy deteriorates in a systematic and continuous manner through the parafoveal and peripheral regions. A viewer’s eye movements reflect the difficulty of the task (Rayner, 2009). During a search task, both the number of fixations and fixation durations (i.e. the length of fixations) increase as the task becomes more complex. During a particular fixation, the area around which search for targets can be conducted effectively is referred to as the perceptual span, or span of effective vision. In relation to the perceptual span of visual search, Rayner and Fisher (1987) demonstrated that it’s size varied as a function of the difficulty of discriminating targets from distractors. This is important as it implies that the area over which effective search can be conducted can be small if the task is difficult.

the number of fixations and the set size of search. Additionally, Findlay (1997) and Findlay, Brown, and Gilchrist (2001) found evidence in support of Duncan and Humphreys' (1989) Resemblance theory reporting that eye movements were directed towards objects that shared features with the target object, while others have found that the number and length of fixations increases when distractors are similar to targets (Rayner & Fisher, 1987). Wolfe (2007) pointed out that the imposed seriality of eye movements supports a serial stage in guided search, however multiple items may be processed in parallel on each fixation.

Using measures of response accuracy, speed and eye movements, data from visual search studies have revealed a number of influences on efficiency of search. A number of models have been applied to account for the data, and GS4 currently represents the most recent developments in the debate.

Searching in Scenes

The accounts of visual search described above are designed to explain data from laboratory experiments where targets and distractors are clearly differentiated from each other and typically appear on a neutral uniform background. Scenes are much more variable and also have semantic content.

A scene is described as a real world environment viewed from a particular vantage point (Henderson, 2011). Typically, a scene will be comprised of background elements (i.e. larger-scale, immovable surfaces and structures) and objects (i.e. multiple smaller-scale elements that can be manipulated) arranged in a semantically coherent way (Henderson & Hollingworth, 1999). However, such a definition depends on spatial scale as scenes can encompass a hierarchical spatial structure in that smaller-scale scenes can form parts of larger-scale scenes, for example, a desk could be considered a scene within a larger scene of an office (Henderson, 2011).

Early studies examining how people look at pictures found that viewers concentrate on interesting and informative regions of a scene (Buswell, 1935; Yarbus, 1967). Furthermore, early evidence was found for task influence as Yarbus (1967) found that giving participants different instructions while viewing a picture produced different eye movement scanning patterns for that scene. The influence of either bottom-up stimulus based information or top-down knowledge based information has been a focus of interest in research examining how we percieve scenes, to establish exactly what constitutes an interesting region in a picture and what drives search.

Three general approaches have been used to assess the influence of bottom-up stimulus based information in scenes (Henderson, 2011). One, a 'scene statistics' approach compares areas of the image that receive fixations compared to a control set of locations that do not. This approach has revealed that fixations occur on regions of higher edge density (Mannan, Ruddick, & Wooding, 1996) in addition to regions of high contrast (Reinagel & Zador, 1999).

A second approach involves building computational models that aim to predict fixation locations. A principle model being the 'saliency-map' approach (Itti & Koch, 2000, 2001), which reveals regions of a picture that differ from their surroundings based on visual properties such as colour, contrast or edges. Regions with uniform features are thought to be uninformative (Treisman & Gelade, 1980) therefore regions that differ from their surroundings are likely to be informative to the viewer (Henderson, 2011). It has been demonstrated that the salient points of the map can be correlated with a viewer's fixations (Oliva, Torralba, Castelhano, & Henderson, 2003; Parkhurst, Law, & Niebur, 2002). However, a number of questions have been raised in relation to the saliency map approach, including whether a single map is retained over multiple fixations or whether new maps are generated, which specific image properties should be considered in the map and whether or not saliency maps can be established as biologically and psychophysically plausible accounts (Henderson, 2011).

The scene statistics and the saliency map approaches are both limited as explanations due to depending on correlational methods, which do not reveal a causal link. For example, although it has been demonstrated that there is a link, or relationship, between salient regions of an image and a viewer's fixations, it is unknown, however, whether the salience necessarily causes the fixations or not, just that they coincide. The third approach to assess the influence of bottom-up information involves manipulation of the scene to establish a causal link. The number of studies directly manipulating image features is low (Henderson, 2011), but Mannan et al. (1995) found that filtering scenes of mid and high-level spatial frequency information (i.e. producing a blurred version of the image) did not change the locations participants fixated, suggesting that viewing is led by image features rather than the need to identify specific objects.

All three of these bottom-up stimulus-based approaches focus on image properties, potentially at the expense of top-down knowledge based factors. In fact, Foulsham and Underwood (2011) suggested that salient regions may coincidentally also be knowledge based informative regions of the scene, useful for interpreting the image or carrying out the search task required. Furthermore, other studies have found that if a blank area of a scene becomes task-relevant, viewers will fixate it (Ferreira, Apel, & Henderson, 2008), contradicting the notion that image features alone guide fixations. Others have suggested that the saliency of a scene will be overridden when a scene contains meaningful information, supporting evidence of top-down knowledge influencing search (Henderson, Weeks Jr, & Hollingworth, 1999; Oliva et al., 2003). The definition of potentially informative areas within a scene, determined on the basis of a rapid initial glance at a scene, is referred to as scene gist (Biederman, 1972; Castelhano & Henderson, 2007). Neider and Zelinsky (2006) found that fixations were frequently drawn to an area in which one would expect to find a target in a scene e.g. finding a hot air balloon in the sky.

Henderson and Ferreira (2004) categorised the different sources of top-down influence on scene perception. First, short and long-term episodic scene knowledge can inform the viewer about where an item has recently been placed, or where said item typically resides, incorporating memory retained over time. Second, scene schema knowledge provides generic semantic and spatial knowledge about a certain category or type of scene, to indicate where something is likely to be (Biederman, Mezzanotte, & Rabinowitz, 1982; Henderson et al., 1999; Oliva et al., 2003). Lastly, task knowledge guides search based on the goal of the viewer (Castelhano, Mack, & Henderson, 2016; Henderson, Malcolm, & Schandl, 2009; Yarbus, 1967).

Searching in the Real World

The studies of visual search in scenes discussed so far do not always represent the complexities and uncertainties of real-world search tasks. Often, these tasks can be safety critical, and/or require high levels of accuracy. Two real world search tasks that have been studied extensively in terms of applied visual search are medical professionals searching images of the human body for abnormalities and airport luggage screeners searching images of luggage for threatening items.

Medical Professionals

Within the medical domain there is a growing reliance on imaging equipment (Reingold & Sheridan, 2011), and radiologists must strive for efficient, accurate and exhaustive search. Compared to groups of inexperienced novice searchers and medical students still in training, experts in radiology demonstrate more efficient scan paths (Krupinski et al., 2006; Krupinski, 1996), faster times to fixate an abnormality (Krupinski, 1996; Kundel, Nodine, Conant, & Weinstein, 2007; Nodine et al., 1996; Nodine, Mello-Thoms, Kundel, & Weinstein, 2002) and a greater proportion of time spent fixating abnormalities in the image (Kundel & Nodine, 2004).

Birkelo et al. (1947) sought to discover the most appropriate viewing method for detecting tuberculosis in chest images. He found that there was a surprising degree of both inter- and intra-observer variation in detecting tuberculosis from x-rays. Further research confirmed this basic finding (Krupinski, 2010; Manning, Gale & Krupinski, 2005). This was even true in simple tasks such as being asked to describe the physical characteristics of radiographic shadows (Newell, Chamberlain, & Rigler, 1954). From investigating the variation in diagnostic decisions grew an interest in understanding the types of errors made by radiologists, what causes the errors to occur, how often they occur, and, most importantly, how the errors could be prevented (Krupinski, 2000).

Unlike traditional visual search tasks, where the objects being examined are unambiguous and easy to identify, medical images contain ambiguities (Krupinski, 2010). Despite these ambiguities, most radiographers agree on the sites of potential interest for diagnosis. Krupinski et al. (2006) examined what attracts fixations during medical image search. The issue of interest was whether fixations are made to regions of diagnostic importance. Three pathologists, three pathology residents and three medical students had their eye movements recorded whilst searching slides from a set of twenty breast-core biopsy surgical pathology cases. The participants were instructed to choose three locations on the image that they would most like to zoom in on if they were going to view the image in more detail, in order to make any diagnostic decisions. It was found that there was a high level of consistency in the areas that were deemed to be of importance or interest across groups (Krupinski et al., 2006). A senior pathologist confirmed that ninety-two per cent of the common locations contained information applicable for diagnosis.

Despite agreeing on the sites of potential interest for diagnosis, radiologists are susceptible to interpretation errors. Interpretation errors come in two kinds: false positive errors (i.e., detecting a tumour that is not present) and false negatives (i.e., failing to detect a tumour altogether). The consequences of false negative errors and missing targets could have serious implications. To begin to understand how false negatives occur, Tuddenham and Calvert (1961), had a group of radiologists use a spotlight while they searched radiographic images. They concluded that false negatives were the result of incomplete image searches. Kundel, Nodine and Carmody (1978) analysed false negative errors and classified them into three categories on the basis of fixations and dwelling times, i.e. how long items were inspected for. The first category comprised search errors where lesions were never actually fixated. The second category consisted of recognition errors, where lesions were looked at but not for a long enough period of time to be able to accurately identify them. The third category consisted of decision errors, where radiologists fixated lesions for a long period of time (above the recommended threshold for detection), but either did not recognise lesions or actively dismissed them (Krupinski, 2010). This classification of false negative error types has now been used when examining lung images (Kundel, Nodine, & Krupinski, 1989), bone images (Hu, Kundel, Nodine, Krupinski, & Toto, 1994), and mammograms (Nodine et al., 2002) as well as being adapted for other search tasks (Cain, Adamo, & Mitroff, 2013; Godwin, Menneer, Riggs, Cave, & Donnelly, 2015).

In addition to the three basic classes of error described above, the study of medical professionals brought to light another class of error. Success in basic visual search tasks is often thought to occur when single targets are detected. With the detection of targets in medical images (and other forms of real world search), however, more than one single target may be present in each instance. When multiple targets are present, but search is terminated before all possible targets are found, then this forms a fourth class of error referred to as inappropriate satisfaction of search (Tuddenham, 1962). For example, if a radiologist is searching an image and detects one lesion or abnormality, this discovery provides 'satisfaction' that they have searched the display exhaustively and they therefore terminate their search prematurely (Tuddenham, 1962). For this reason, lesions can be missed if accompanied by others, which are detected first (Berbaum et al., 1990). Although estimates regarding the importance of satisfaction of search differ, Berbaum et al. (2010) found that satisfaction of search could account for up to one third of misses in radiology. Others have revealed that in some cases, searchers continue to search after finding a first target, but still miss additional targets (Berbaum, Franken, Dorfman, & Caldwell, 2000). This suggests that the error may occur for reasons other than 'satisfaction' (see next section) and consequently there have been suggestions to rebrand the phenomena the 'subsequent search misses effect' to avoid confusion (Adamo, Cain, & Mitroff, 2013). However branded, the risk of missing an additional target once an initial target has already been located could have important implications in many real world search tasks in which multiple targets are present.

Airport Security Screeners

An airport security screener is typically required to search an x-ray image of a bag or suitcase for threatening items such as knives, guns and Improvised Explosive Devices (IEDs). The incidence of real IEDs in baggage is very low indeed. In a typical laboratory visual search task, the target is presented on 50% of the trials. Recent research has begun to explore the influence that target prevalence (i.e., the proportion of trials where a target is presented) has upon search performance. What happens to accuracy and response times if an observer has to search for a target that appears at a prevalence of less than 50% of the total number of trials? Studies show that as target prevalence decreases, the probability a target will be identified also decreases (Wolfe, Horowitz, & Kenner, 2005). As a result, the 'prevalence effect' occurs following the rare presentation of targets throughout a succession of trials, causing targets to be neglected once finally presented (Wolfe et al., 2007, 2005).

Early work by Wolfe et al. (2005) employed an artificial baggage screening task to compare search performance between prevalence levels of 1%, 10%, and 50%. Observers missed only 7% of the targets when target prevalence was high (50% prevalence), but this increased to missing 30% of the targets when the target prevalence was low (1% prevalence). Fleck and Mitroff (2007) hypothesized that the higher miss rates in low relative to high prevalence search conditions arises from motor priming (e.g. the observer detects the target but responds too quickly to give an accurate response). Adopting a similar experimental design to Wolfe et al. (2005), they gave participants the opportunity to correct a previously given response, with the belief that participants could correct errors caused by motor priming but not those caused by an error in perception. Fleck and Mitroff (2007) found that the opportunity for correction removed the prevalence effect. They argued that the prevalence effect was the result of participants simply responding too rapidly, despite having correctly detected the target.

Further research questioned this account and it was discounted as a complete explanation, following a number of studies in which motor priming was controlled for (Godwin et al., 2010, 2015; Wolfe et al., 2007). Wolfe et al. (2007) suggested that the prevalence effect was due to a shift in the response criterion, rather than a motor priming speed-accuracy trade off. When paired participants were presented with the same set of stimuli, it was found that errors were highly correlated between the pair; in other words, it was likely that two participants, shown the same stimuli, missed the same targets. Wolfe et al. (2007) concluded that if the prevalence effect occurred because of rapid, careless responses, then there should not have been such a high correlation in the targets that were missed. In other words, Fleck and Mitroff (2007)'s account of the prevalence effect was unlikely to account for these results.

Others have explored the prevalence effect using a relative prevalence paradigm. Godwin et al. (2010) used a multiple target search task in order to manipulate the prevalence of different target objects, but to keep the overall target probability constant. This meant that motor priming was controlled for and decisions to terminate the trial early were avoided (i.e. the addition of a second higher prevalence target meant that searchers could not terminate search prematurely). They found that higher prevalence targets were better detected than lower prevalence targets. Even when the overall target prevalence indicates that targets will occur reasonably often, a specific low probability target is still likely to be missed. Godwin et al. (2015) extended this and found that although participants were equally likely to fixate both low and high prevalence targets, they tended to fail to identify low prevalence targets after fixating them, suggesting that prevalence effects were due to failures to identify the target, after fixating it. Supporting this, Hout, Walenchok, Goldinger, and Wolfe (2015) also found that even when participants directly look at the low prevalent target, they often fail to detect it.

The prevalence effect has more recently been thought of as emerging from two decisions, one perceptual and one cognitive (Wolfe & Van Wert, 2010; Wolfe et al., 2007). The first decision occurs at a perceptual level of deciding whether the individual object currently being examined is a target or not. The second occurs at a cognitive level, concerning the search array as a whole, and whether or not the target is found before the threshold is reached to cause searchers to terminate search (Chun & Wolfe, 1996; Schwark, MacDonald, Sandry, & Dolgov, 2013). During low prevalence search, this threshold shifts as participants make rapid target-absent responses allowing themselves only a short length of time to detect the target before they terminate search.

The ‘satisfaction of search’ (Tuddenham, 1962) error or ‘subsequent search misses effect’ (Adamo et al., 2013) has been demonstrated in the context of airport security as well as in medical imaging (see previous section). A series of studies carried out by Mitroff and colleagues have built upon early findings of satisfaction of search and explored misses in multiple target search tasks (Adamo et al., 2013; Biggs, Adamo, Dowd, & Mitroff, 2015; Cain et al., 2013; Cain, Biggs, Darling, & Mitroff, 2014; Cain & Mitroff, 2013; Clark, Cain, Adcock, & Mitroff, 2014; Mitroff & Biggs, 2014). Cain and Mitroff (2013) examined two different theories thought to cause interference in multiple-target search; the perceptual-set theory and the resource depletion theory. The perceptual-set theory suggests that after one target is found (e.g. a pistol in a luggage x-ray), searchers are prone to focus search for other targets that are perceptually similar to the first one (e.g. another gun) and are therefore less likely to detect targets that are perceptually dissimilar to the first target (e.g., an IED). The resource depletion theory suggests that after one target is found, its identity and location is stored in working memory, which depletes cognitive resources which in turn causes errors in detecting a second target (Cain & Mitroff, 2013). Using a search task requiring participants to search for T-shapes amongst L-shapes, Cain and Mitroff (2013) found support for the theory of resource depletion. Replacing found targets in the display with a random distractor did not improve subsequent search, suggesting subsequent search miss errors are due to something more than just perceptual salience of the previously found target. However, removing the found target or making it salient and easily segregated from the other items improved subsequent search. Cain and Mitroff (2013) concluded that working memory load has a larger effect on accuracy for subsequent targets than salience.

Cain et al. (2014) suggested that this working memory deficit could be improved. They demonstrated, in both undergraduate participants and professional visual searchers from the Transportation Security Administration, that dividing a multiple-target search into several single-target searches (separated by three to five unrelated trials) effectively freed the working memory resources used by the found target and eliminated the subsequent search miss errors. However, more recently, Biggs et al., (2015) have used a mobile application to collect a large set of data from an 'Airport Scanner' game, in which the player acts as a airport security officer and searches for 'illegal' targets in simulated x-ray images of baggage. Biggs et al., (2015) found that subsequent search miss errors were reduced when the two targets in the multiple-target search display were identical, suggesting support for the perceptual-set theory. Further analysis revealed that this perceptual bias could come in two forms, either perceptual-set bias (i.e. interference relating to the features of the first target) or conceptual-set bias (i.e. interference relating to the conceptual relationship of the first target). Others have also found subsequent search miss errors to be accounted for by attentional-blink like effects (Adamo et al., 2013). It may be that subsequent search miss errors have multiple causes (Cain et al., 2013).

Recent work has also explored what factors influence search for multiple targets in order to decipher ways in which search efficiency can be improved. A number of situational factors have been found to increase the negative effects of multiple target search (Adamo, Cain, & Mitroff, 2015; Mitroff, Biggs, & Cain, 2015). For example, compared to single-target search, errors were found to increase in multiple target search when searchers were anxious (Cain, Dunsmoor, LaBar, & Mitroff, 2011), when search was for a fixed number of images, as opposed to for a fixed amount of time (Clark et al., 2014), when a time pressure is introduced (Fleck, Samei, & Mitroff, 2010) when searchers expect a certain number of targets to be present (Cain, Vul, Clark, & Mitroff, 2012; Fleck et al., 2010) and when there is more 'clutter' (i.e. the number of other items close by) around a second target (Adamo et al., 2015).

Mitroff et al. (2015) has made three suggestions to combat the negative influence of multiple target search. These suggestions were made in relation to airport security screening, but could be applied to other domains. 1) Move the screening process to a remote location. This could reduce anxiety, allow bags to be viewed by more than one screener, improve the structure of work flow in terms of staff scheduling and time-based breaks, and allow assessment images to be integrated. It is also important to note here, that in the UK advances are currently being made in introducing 'Centralised Image Projection' into airports. In this, a second searcher checks the performance of an initial searcher, in order to help with efficiency and more challenging cases. Future advances may see this format also moved to a remote location. 2) The number of prohibited items could be reduced (i.e. remove items such as water bottles). This would reduce the number of targets likely to appear, reducing the risk of a subsequent miss error, increase search speed by reducing memory demands for multiple targets (Wolfe, 2012) and allow high threat items to become the priority, regardless of their salience. 3) Implement more training that emphasises the importance of search consistency, teaching screeners to execute their search in the same way from one bag to the next. Systematic search would mean less of a burden on memory (Cain & Mitroff, 2013), increased search speed at the conveyor belt and does not require new technology to be implemented. Mitroff et al. (2015) suggest that further research is needed to fully understand subsequent search misses but hope that their suggestions may be a step towards combatting the negative influence of subsequent search misses in the applied setting of airport security screening.

In addition to the implications of searching for more than one target (i.e. multiple instances of a gun) in a display, compared to searching for a single target alone, searching for two different target *types* (i.e. guns, knives and IEDs simultaneously) leads to a cost in performance, known as the dual-target cost. This cost has primarily been demonstrated in response accuracy, but in addition, has also been demonstrated in time and efficiency of search (Menneer, Barrett, Phillips, Donnelly, & Cave, 2007). The dual-target cost results from a decrease in search guidance to target objects when more than one target might occur relative to when only single targets are presented (Menneer et al., 2012; Stroud, Menneer, Cave, & Donnelly, 2012; Stroud, Menneer, Cave, Donnelly, & Rayner, 2011). There are two ways in which two separate targets may be mentally represented (i.e. the structure of the image that is stored in our mind) so that they can be searched (Menneer, Donnelly, Godwin, & Cave, 2010). First, each possible target might be represented separately. Second, a single coarse representation may be formed and include both the range of features present in each target and all feature values that lie in between the two targets in some feature space (Stroud et al., 2012). A number of studies have utilised the recording of eye movements to explore guidance and the role of mental representations in dual-target search. Menneer et al. (2012), for example, examined guidance in dual-target search while also exploring whether the dual-target cost remains when stimuli (x-ray images) overlap. They replicated the dual-target cost, and demonstrated a reduction in guidance in dual-target search, in terms of fewer fixations to target-colour objects, and more fixations to non-target colour-objects, compared to single-target search.

Stroud et al. (2011) used stimuli comprised of abstract shapes modelled after threat items to explore the role of colour in guiding dual-target search. Eye movement data revealed that when two targets were different colours, more distractors with colours different from either target colour, were fixated more. This was in line with previous work suggesting the dual-target cost is more prevalent in dual-target search for dissimilarly coloured targets compared to two targets that share colour (Menneer, Cave, & Donnelly, 2009). Stroud et al. (2011) suggested that colour information guides search, meaning that searchers are unable to represent two coloured targets simultaneously. Stroud et al. (2012) found a similar result using a coloured T-shape target amongst L-shapes showing that as the similarity between the two target colours decreased, search efficiency decreased. Stroud et al. (2012) concluded that in terms of mental representations, when the two targets are dissimilar, they appear to be encoded as separate representations. However, when two targets are similar the mental representations guiding search can be viewed either as separate representations or as a single range containing the two targets, and the colours between them in space.

The results of Menneer et al. (2012, 2007, 2010) and (Stroud et al., 2012, 2011) point to similar conclusions. If search is for multiple target types, then a dual-target cost will be found, which in turn is likely to lead to some targets being missed. It is also worth noting that Menneer et al. (2009) and Menneer et al. (2012) also found that the dual-target cost held after practice at the task. In the context of security screening, suggestions have been made to avoid the dual-target cost by encouraging specialist screeners to search for specific categories of threat only (Menneer et al., 2012, 2007, 2009). For example, screeners could use colour to guide their search by having one searcher look for knives and guns (which are blue in colour on the x-ray images) and another IEDs (which are orange in colour on the x-ray images) to ensure efficient search (Stroud et al., 2012, 2011). Menneer et al. (2012) estimated that the addition of a second specialised screener would reduce the number of missed targets by 28% and furthermore, increase the speed of passengers passing through security by 68%. Alternatively, screeners could be trained to conduct two separate single-target searches of the same display, i.e. search for guns first and then for IEDs (Menneer et al., 2012), which previous research has suggested may be possible following instruction (Beck, Hollingworth, & Luck, 2012).

Exploring search in real world scenarios has revealed a number of challenges. Targets may be missed, but also may not be recognised or may be misinterpreted. The likelihood of the target appearing may be unknown, or (problematically) very low. Furthermore, the searcher may be faced with the task of searching for more than one type of target at the same time, or may be required to detect multiple instances of the same target type appearing in one display. These issues build the uncertainty associated with real world search, and make it difficult for the searcher to know when to stop searching. When targets appear infrequently, you may be led to terminate search prematurely. Likewise, finding one target in a multiple target display may lead to early termination of search following this initial find. The question of when to stop searching is an important aspect of many real world search tasks, indeed, how do you know when you've searched for long enough?

Foraging

All the studies discussed so far have considered search as a task completed in relation to single images (including sets of images). However, in some tasks search must stretch across different instances or environments. In these tasks a further complication to search is introduced in deciding when one search should end and another begin. Uncertainties around target identity and prevalence, and search exhaustiveness make the decision to move on a difficult one to make. Decision-making becomes more complicated as they become 'trade-offs' between optimising intake and knowing when to move on. It is an issue that has been explored in the context of work on animal foraging. Specifically how target distributions or learning the statistical properties of the environment inform decisions about when an animal should move on to a new patch in search of food (see Stephens & Krebs, 1986, for a review).

Animals forage to find wild food, and it is therefore essential that they are efficient in order to survive. Foraging theories have therefore focused on optimising foraging pay-out. The optimal foraging theory (Charnov, 1976; MacArthur & Pianka, 1966; Pyke, Pulliam, & Charnov, 1977) suggests that if an animal consumes food from a number of locations or 'patches' in an environment, they will strive to maximise their intake of food. To do this, when it comes to deciding to move on from one patch to another, they will move on once the intake for the current patch falls below the anticipated rate for the environment as a whole. Therefore the optimal foraging theory suggests that to maximise overall consumption, the animal will move on from a particular patch before all the food has been foraged i.e. the animal will not complete an 'exhaustive search' for that area.

A number of authors have noted a relationship between visual search and animal foraging. In a series of experiments using a visual search task in which participants were required to 'pick' berries from multiple berry patches, Wolfe (2013) found that humans adapted their search in a similar fashion to animals. Participant's behaviour was rule-governed and could be predicted by the marginal value theory (Charnov, 1976), a dominant concept in optimal foraging theory, which states that time (in terms of the rate of extraction from the current patch and the time to the next patch) is an important determinant in when you will move on your search. Similarly, Cain, Vul, Clark, and Mitroff (2012) found that participants adjusted their search time in visual search for multiple targets. If an additional target was more likely to occur, participants adapted to the statistics of the search environment and searched for longer. In contrast, if an additional target is less likely to occur, then participants would terminate their search more quickly. In this case, the results were predicted by the Optimal Foraging Theory, potential value theorem; In this case, the results were predicted by the Optimal Foraging Theory, potential value theorem; the forager predicts the 'value' of staying or moving from their current location (or trial) based on a continuous track of how much time has been spent searching, and how many targets have been found (McNamara, 1982). Cain et al. (2012) noted, however, that although participants adapted to the target distribution, they did not adjust as much as would be considered 'optimal', which Cain et al. (2012) suggested was due to participants taking a more conservative approach when searching target absent trials.

Optimal foraging theory proposes a simple rule; that searchers will move on when intake for the current patch falls below the average rate of the environment as a whole. However, this assumes that this average for the environment is fixed, and assumes that decisions are made based on an expected rate of intake, ignoring the searchers belief about how many targets may be present, and how easy they may be to find. A number of studies have extended the optimal foraging theory to address these issues.

A recent study by Ehinger and Wolfe (2016) aimed to apply the optimal foraging theory to a more complex task, searching for targets in a natural scene, i.e. petrol stations in a large satellite image. They suggested such a task is more complex than previous foraging studies due to the higher-level knowledge and decision-making processes that coincide with the task. Specifically, these forms of search tasks differ from classic laboratory search tasks in five ways, 1) a scene is continuous, so set size can be impossible to measure, 2) it is hard to know how many items in a scene are processed in a single fixation, 3) the number of targets is unknown, 4) the structure of the scene guides searchers to where targets may be likely to appear and 5) each stimulus is searched for minutes, rather than a fraction of a second. To incorporate these factors, Ehinger and Wolfe (2016) proposed an optimal foraging *model* that shifted focus onto the participants' expected rate of intake, rather than the observed rate of intake. Ehinger and Wolfe (2016) found their model to be a good predictor of searchers' quitting time in the complex search task examined.

Fougnie, Cormiea, Zhang, Alvarez and Wolfe (2015) and Zhang, Gong, Fougnie and Wolfe (2015) argued that previous studies in human optimal foraging assume that the average output of the environment is fixed, and do not consider the relationship between the quality of the immediately preceding patch(es) searched and the quality of the current patch being searched. Fougnie et al. (2015) pointed out that the real world has temporal structure, such as seasons, that may influence foraging e.g. there would likely be more apples to forage during autumn compared to winter. Fougnie et al. (2015) explored this by having participants search for multiple 'T' shapes amongst 'L' shaped distractors as target density either rose or fell systematically across trials. They found that participants searched for longer when the 'quality' of the patch was falling (i.e. decreasing in target density) compared to when it was rising. This was inconsistent with marginal value theorem (Charnov, 1976) of the optimal foraging theory, which would predict that searchers would move on as targets decrease in number. Furthermore, when interupted and asked to rate patch quality, participants gave a higher quality patch rating during falling phases, than during rising phases, even though the actual quality was equivilant at the times they were asked. Fougnie et al. (2015) concluded, therefore, that the history of search alters both participants' behaviour and beliefs, and foraging theories need to consider more detailed models to explain this behaviour. Zhang et al. (2015) also explored this issue in a foraging task modelled on Wolfe's (2013) berry picking task. Participants 'picked berries' through periods of high and low density patch quality. Zhang et al. (2015) found similar results to Fougnie et al. (2015), with participants also searching for longer when the patch quality was falling than when it was rising.

Recent work has also begun to explore what they term 'hybrid foraging', integrating a memory component too. If foraging search is defined as collecting multiple instances of a single target type, and hybrid search is defined as the search for one instance of several types of targets held in memory, then 'hybrid foraging' combines the two and is defined as the search for multiple instances of any type of several types of targets. Wolfe, Aizenman, Boettcher and Cain (2015) found that in this type of search task, participants tended to collect items in 'runs' of one target type at a time. Wolfe et al. (2015) suggested that this leads to more efficient search times, as once an initial target is selected, the next target to be selected is guided by the search for the previous target, creating a bias towards the first target type. A similar result was found by Kristjánsson, Jóhannesson and Thornton (2014), who suggested that typically, when food items are conspicuous, an animal may switch between available food types at random, but when items are hard to detect, animals are more likely to focus on a single food type. They replicated this in human search using an Ipad task in which participants located and touched 40 targets from *two* categories. Kristjánsson et al. (2014) found that when target items 'popped out' and were distinguishable by colour, participants frequently switched between categories.

However, as soon as focused attention was required to identify targets, participants exhausted one category before searching for the other.

These studies have added to the growing literature replicating animal foraging behaviour in humans. Two other aspects of animal behaviour worth considering here are Lévy flights and area-restricted search. Area-restricted search suggests that when animals detect areas with a higher density of resources, they will make a higher number of turns. This acts to localize them to where valuable resources are most dense. Kalff, Hills and Wiener (2010) replicated this in human searchers using a large virtual environment, in a task in which participants collected mushrooms in a meadow. When the targets were scattered in patches (as opposed to evenly distributed) participants increased their turning rate and turned more sharply after encountering high density patches of resources. Lévy flights are random walks in which movement lengths are independently drawn from a probability distribution, i.e. underlying movement distributions to not change in response to resource encounters. Hills, Kalff and Wiener (2013) suggested that in patchy environments with areas of higher density resources, humans searchers show evidence of area-restricted search, which is likely to be more efficient than a Lévy flight. However, Lévy flights can be utilised in spatially uncorrelated environments.

The studies above suggest that humans can be sensitive to the statistics of the environment. While this sensitivity may be helpful in some scenarios, as it is with animals foraging for food, there are other scenarios when it is unlikely to be helpful, for example, when required to conduct an 'exhaustive' search of an environment, i.e. search should only be terminated once all locations have been searched, rather than when a target(s) has been detected. A number of recent studies have found that in search tasks requiring the searcher to be careful and exhaustive, experience can lead to an increasingly cautious approach, with increased efficiency emerging from a slower, more thorough search strategy. For example, Jackson et al. (2013) found that orthodontists (who were specifically trained to judge facial symmetry) were slower but more accurate than non-trained students at assessing facial symmetry in both upright and inverted faces.

Likewise, when examining visual performance in airport security luggage screeners, Biggs, Cain, Clark, Darling, and Mitroff (2013) found that compared to students the airport screeners were more accurate, took less time to locate a target, and critically, spent longer searching the display before terminating their search. Biggs et al. (2013) also found that consistency (in terms of similar individual search times from trial to trial) was higher for the airport screeners particularly for those who were more advanced in their career. Biggs et al. (2013) suggested that the consistency in prolonged search times was a learnt search strategy. These findings were demonstrated in a classic visual search task (find the ‘T’ shape amongst pseudo ‘L’ shape distractors), which ensured that the airport screeners did not hold an advantage in terms of practice or familiarity with targets, which may have been the case if the task was to search an x-ray image for a gun or a bomb. Extending this, Biggs and Mitroff (2013), using a similar design, found that in comparison to the nonprofessional searchers, the airport screeners took longer to find a first target, and took longer to discontinue their search after finding this first target. It could be argued that the airport screeners had a greater incentive to perform well in the task, and were therefore more motivated. To address this, within the airport screeners Biggs and Mitroff (2013), compared early career and experienced professional searchers, who were matched on general motivation. No differences were found for speed or accuracy overall, however, consistency in response times between trials was found to be the best predictor of accuracy for the experienced professional searchers, whereas speed was found to be the best predictor of accuracy for the early career professional searchers. Biggs and Mitroff (2013) argued that this difference between the professional searchers lessens concerns regarding motivation. Overall, Biggs et al. (2013) and Biggs and Mitroff (2013) reveal an important finding, which indicates that training and experience may help combat the early termination of search problem found to occur in a number of search scenarios as discussed previously.

It is also interesting to note that Jackson et al. (2013) also tested airport security luggage screeners in their facial symmetry task. Although the airport screeners did not succeed in the increased accuracy that the orthodontists achieved (in fact their accuracy performance was equivalent to that of the students) their search times were comparative to those of the orthodontists. Together, the results from Jackson et al. (2013), Biggs et al. (2013) and Biggs and Mitroff (2013) suggest that to improve performance, training in some types of search tasks actually slows people down and leads to a more conservative search strategy. Additionally, this search strategy can be extended into other search tasks outside the searchers' trained domain.

Movement and Search of Large-Scale Environments

Recent research has begun to explore search tasks taking place on a larger-scale, incorporating eye, head, and even body movements as searchers physically explore the environment. A number of studies have shown that searching large-scale environments is rather different from visually searching on computer screens.

To explore whether the typical properties of a standard visual search task apply to a large-scale physical search task, Gilchrist, North and Hood (2001) had participants physically search through an array of film canisters which were displayed in a grid on the floor of a room, for a marble target. They found that search time increased linearly with set size (the number of film canisters on the grid) and the search slope for the target present display was half that of the target absent slope, consistent with a serial self-terminating or 'inefficient' visual search (Wolfe, 1998). Where the results differed was in the tendency for participants to recheck distractor locations. In contrast to standard visual search, rechecking searched locations was rare with participants making fewer revisits to canisters they had already searched. Gilchrist et al. (2001) suggested that there is a higher cost (in terms of the physical effort to cross the room) associated with forgetting searched locations in large-scale search tasks and therefore, perhaps the increased effort required to search leads to a higher likelihood of participants remembering searched locations. A similar result was reported by Ruddle and Lessels (2006), who found that participants made less revisits when they were required to walk through a virtual display as opposed to viewing a virtual display screen. Furthermore, Solman and Kingstone (2014) used a novel head-contingent display system to directly compare search using head movements and search using eye movements. They found an increased use of memory in head-contingent compared to eye-contingent search, as search became more effortful. This was demonstrated in longer initiation times before inspection, suggesting more resources were devoted to recall efforts prior to initiating perceptual search. Head-contingent searchers were also more likely to orient directly to the target items, and were faster to respond once the target was detected.

As with the real world search tasks performed on computer screens and discussed previously, targets in many large-scale real world search tasks are unlikely to be plainly visible. Smith, Hood and Gilchrist (2008) extended Gilchrist et al.'s (2001) findings by comparing visually guided and non-visually guided search in the same large-scale experimental setting, using an array of lights and switches displayed on the floor. In the visually guided condition, the colour of the light (i.e. a single red light amongst red and green combined lights) guided the participant to a target switch, representing a typical visual search task. In the non-visually guided condition, the participant had to physically inspect each location as the target light only appeared once the switch had been pressed, then representing a similar experimental set up to Gilchrist et al. (2001). As expected, Smith et al. (2008) found that results in the visually guided condition were comparable to a feature visual search task (Treisman & Gelade, 1980) in which the targets 'popped out'. In the non-visually guided condition, Smith et al. (2008) found similar results to Gilchrist et al. (2001) in that search was conducting in a serial self-terminating fashion, due to the need to physically inspect each location.

In another large-scale search task, Jiang, Won, Swallow and Mussack (2014) required participants to search for a coin in a large outdoor area and report which side of the coin was facing up. Jiang et al. (2014) found that by manipulating the likelihood of the coin appearing in a particular quadrant of the display, participants showed evidence of statistical learning effects and were sensitive to environmental cues, suggesting that environmental cues may play a more crucial role in large-scale search. Smith et al. (2008) also found that when the target light was more likely to appear on a certain side of the room, participants were more likely to detect the target in that location, supporting Jiang et al. (2014)'s finding. Furthermore, Won, Lee and Jiang (2015) extended on work by Jiang et al. (2014) and Smith et al. (2008) to address whether these effects occur in early spatial attention, or later response decisions, i.e. do participants become faster to orient or do they become faster to respond to targets in high probability locations. Also requiring participants to detect a coin in a large-scale task, Won et al. (2015) replicated Jiang et al. (2014)'s finding that participants were faster to detect coins in the high probability quadrant and furthermore, by measuring head movements, found that participants' head movements were highly sensitive to the location probability of the coin. Won et al.

(2015) concluded that learning effects influence attentional orienting early on during search.

Foulsham, Chapman, Nasiopoulos and Kingstone (2014) also used a large-scale search task and had participants search for a pre-defined mailbox in a University mailroom. They manipulated the saliency of the mailbox, and the priming to the saliency of the mailbox. They found that the saliency of the mailbox only influenced performance when participants were primed to its existence, while salience is a cue known to be important in visual search (Itti & Koch, 2000, 2001).

Searching in Pairs or Teams

All of the research reviewed so far has concerned individuals working alone to find targets. However, in many real world search tasks, particularly threat detection, individuals work as part of a team. Accounts of search should, therefore, take this into account. Individuals work in teams because it is believed that collaborating and sharing with others can improve efficiency (Goldstone, Roberts, & Gureckis, 2008). However, team working (as distinct from individuals working alone) is known to be subject to multiple social influences. Attempts to share workload can lead to social loafing unless all members of the team cohere and have a shared view of the importance of completing the task proficiently and successfully (see Karau & Williams, 1993 for a review). Also, more helpful, positively oriented and extraverted individuals with high self-esteem, will exhibit social facilitation in joint tasks, whereby performance improves in the presence of others (Strauss, 2002). For more information on the motivation behind social processes, how people interact with each other and some of the social factors that can influence team work, see Hogg and Vaughan (2013). The focus of this review, however, is the influence working in pairs or teams has on the efficiency of search behaviour. There are, however, only a small number of studies which have explored whether 'two heads are better than one' in the context of visual search.

Brennan, Chen, Dickinson, Neider, and Zelinsky, (2008), examined the kind of information that might facilitate or harm collaborative search relative to an individual baseline. Pairs performed the task (detection of a latter 'O' amongst 'Q's) coordinating using either shared gaze (a gaze-cursor indicated where the other searcher was looking), shared voice (the pair were able to speak to each other), or shared gaze and voice. Brennan et al. (2008) found that pairs in the shared gaze and shared voice conditions were quicker (and as accurate) at detecting and identifying targets than participants performing the task alone. They also found participants created virtual boundaries dividing displays either horizontally or vertically. Chen (2007) found that participants would only use spatial division when the task did not allow for more meaningful divisions of labour (e.g. target type). Furthermore, Brennan and Enns (2015) explored whether collaborative benefits derive from the statistical facilitation of independent responses (i.e. the expected benefit of aggregating independent responses), or from social interaction (i.e. efficient collaboration). Using a visual enumeration task (indicate the number of targets present in a display) they also found that two searchers were better than one on a visual search task and concluded that social interaction was the underlying cause.

Summary and Direction for Present Thesis

The goal of the present thesis is to explore factors influencing the effectiveness of search behaviour in complex real world tasks. Chapter 1 has considered the basic models of visual search, which reveal important findings in relation to when more or less distractors are present, and how features such as colour, shape and orientation make search more or less efficient. Studies exploring how we view scenes and photographs have added to this, revealing what happens when search is guided by semantic content and context. However, real world search tasks are more complex and uncertain than basic stimuli or photographs on a computer display. A number of factors have been discussed as having an influence over efficiency of search in real world tasks, from differences in the types and quantity of targets sought, to the biases and decisions associated with more complex search tasks. These factors need to be considered when providing an account of searching in the real world. Furthermore, the present thesis considers these factors in the context of trying to understand individuals searching for targets alone and when searching with others.

Serching alone or with a partner will be explored across three real world search task scenarios; 1) in the context of airport X-ray security screening on a computer display, 2) in the context of searching for clues to a crime or IEDs in patrol on a large area of grassland and 3) in the context of rummage searching both visually and physically for threatening items in a large-scale house environment.

Chapter 2

Airport Security Screening: How Visual Search Behaviour Adapts Based on Target and Partner Expectations

Some of the data from this chapter has been previously published in Psychonomic Bulletin & Review (Godwin, Menneer, Riggs, Taunton, Cave, & Donnelly, 2016).

In terms of the complex search task of airport security screening, Chapter 1 revealed a number of factors to be important. Of particular interest to Chapter 2 was that search can be incomplete when searching for two target types simultaneously, i.e. guns and IEDs (the dual-target cost; Menneer et al., 2007) and when searching for infrequently occurring targets, i.e. a rare IED (the prevalence effect; Godwin et al., 2010; 2015; Wolfe et al., 2007; 2005). Chapter 2 took these two factors and explored them in relation to how working in a pair might influence search in a task comparable to airport security screening.

In a visual search task, searching for more than one target type simultaneously (e.g., guns, knives and Improvised Explosive Devices in an airport security X-ray baggage search), in comparison to searching for one target type, has been found to lead to a cost in performance, known as the dual-target cost. This cost has primarily been demonstrated in response accuracy, but in addition, has also been demonstrated in time and efficiency of search (Menneer et al., 2007). The impact of this cost is very important for real world safety-critical tasks such as airport security. In this context, suggestions have been made to avoid the dual-target cost by encouraging specialist screeners to search for specific categories of threat only (Menneer et al., 2007, 2012; Menneer, Cave, & Donnelly, 2009). Furthermore, airports in the UK are starting to introduce "Centralised Image Projection", whereby a second searcher checks the performance of an initial searcher. If either of these methods are going to benefit search then the implications of more than one searcher working together as part of a team must be addressed: at a basic level, one would expect search performance to improve when the task of searching for targets is divided, but whether or not this is actually the case needs to be addressed. Chapter 2 explores this issue.

When pairs of searchers are assigned to search for different targets within the same displays then, by definition, the targets must be distinguishably different. Moreover, the targets can appear at differing frequencies both between targets and within the search task. When you are searching with a partner, expectancies can build surrounding your partner's performance, and whether that impacts your own. For example, if you are rewarded or punished as a team for your performance, then, at a basic level, you will seek for your combined performance to be as effective or as efficient as possible. Put another way, the mere act of engaging in a paired search task may have a fundamental shift on your own performance, and how you approach a search task. With that in mind, before engaging in search tasks where participants actually search in pairs, it is important to establish a baseline for understanding the underlying influences that paired search has upon behaviour. In order to establish a baseline for how expectancies can influence search behaviour, this chapter begins with a study focusing on individual rather than paired search. Experiment 1 taps into the expectancies that develop when participants search for targets at different prevalence levels. Experiment 2 taps into the expectancies that develop when participants search with a partner (who was not actually present) who performs at differing levels. To that extent, both experiments focus on how expectancies and predictions both within the structure of the task (prevalence effects in Experiment 1) and within the performance of a partner (Experiment 2) influence behaviour.

Experiment 1

Experiment 1 explores expectancies surrounding the prevalence effect (Wolfe, et al., 2005) As noted in the Literature Review (see Chapter 1), in a typical laboratory visual search task, targets are presented on half of the trials, but this does not always reflect the frequency typically experienced in the real world, where targets often occur infrequently (e.g. in tasks such as airport security screeners searching luggage for threats). Research has demonstrated that as target frequency (*prevalence)* decreases, the probability of target identification also decreases and "target-absent" responses are made more quickly (Wolfe et al., 2005).

A number of explanations have been put forward to explain why target prevalence influences visual search behaviour. Initially, it was thought that the prevalence effect arose due to motor priming of target-absent responses (i.e.

participants adopted to frequently responding 'target-absent'; Fleck & Mitroff, 2007). However, this account has been discounted following a number of studies in which motor priming was controlled for, but prevalence still influenced responses (e.g. Godwin et al., 2010; 2015; Wolfe et al., 2007).

More recent accounts have suggested that the prevalence effect is due to an influence on the participant's response bias, leading to 'fast-absent' responses and target misses in low prevalence conditions (Chun & Wolfe, 1996; Godwin et al., 2010, 2015; Schwark, MacDonald, Sandry, & Dolgov, 2013; Wolfe & Van Wert, 2010). Eye tracking studies have revealed that this response bias effect is demonstrated during the identification of targets, rather than during selection of targets, i.e. participants fixate low prevalence targets but fail to detect them (Godwin et al., 2015; Hout et al., 2015). Experiment 1 explores the source of the bias that leads to the prevalence effect.

Over the course of a study, bias may come from two sources. Target repetition leads to bottom-up priming, and the experience of the frequency of seeing targets can inform top-down expectations about the future likelihood of targets. Previous studies have explored the contribution of top-down priming in the context of target prevalence by manipulating expectations about target prevalence. Lau and Huang (2010) created a situation in which prior and future target prevalence differed (i.e. targets were frequent in the past, but expected to be rare in the future). In addition, they used a coloured frame around visual displays to indicate whether targets were likely or not. Regardless of whether the cue was valid or not, Lau and Huang (2010) showed that target detection accuracy was influenced only by the actual prevalence of the target. In contrast, the time taken to decide that a target was absent was influenced by the cue identity. This suggested that, for target accuracy, the prevalence effect may not be modulated by top-down control by means of expectation (the cues), but results from past experience through repetition of the target. Search termination times, however, may be influenced by cues. A similar result was also reported by Ishibashi, Kita and Wolfe (2012).

Schwark et al. (2013) modulated expectations in a slightly different way, by providing feedback *following* a trial, but where the feedback could falsely report participant accuracy. In their task, participants detected an "X" amongst an array of other letters, responding by either localising the actual target on the display or by just clicking a target-present mouse button. Their results showed target-present mouse clicks were sometimes used to respond when the targets weren't actually present at all, and that localisation responses were used less frequently as set size increased and when target prevalence was high. The results suggested participants made decisions partially based on their expectations.

In sum, the existing data show a mixed pattern. There is some evidence for a role of both bottom-up and top-down influences on the bias that drives the target prevalence effect. Experiment 1 explores the issue of bottom-up information influencing the prevalence effect further. In Experiment 1a, participants searched for a T amongst Ls in trials which alternated between high (95%) and low (5%) prevalence 'slides' presented in pairs. In each pair, the slide order (first or second) allowed participants to form expectations about target prevalence. Participants performed in either an alternating-colour condition or a same-colour condition. In the alternating-colour condition, the target was a different colour for each slide (fixed across every trial) and in the same-colour condition the target colour was the same for both slides of each trial (and again fixed across the experiment). In the same-colour condition, target colour was the same across high and low prevalence slides. In the alternating-colour condition, target colour changed across high and low prevalence slides. It was predicted that changing target colour across slides should lessen bottom-up priming and hence increase the effect of target prevalence on target accuracy and reaction times. As both alternating and same-colour conditions share a top-down component, Experiment 1 is neutral with respect to showing evidence of top-down expectations on visual search.

Experiment 1a

Method

Participants

Sixteen undergraduate participants from the University of Southampton (4 males and 12 females, mean age = 22 years, SD = 4.59) with normal or corrected-to-normal colour vision took part in the study for either course credit or payment. The study was performed in accordance with the Declaration of Helsinki and was approved by the University of Southampton, School of Psychology ethics committee; informed consent was obtained from all participants.

Apparatus

The experiment was produced using SR-Research Experiment Builder. Participants viewed stimuli on a 21” ViewSonic P227f CRT monitor with a resolution of 1024x768 pixels and a refresh rate of 100 Hz. Eye movements were recorded using an Eyelink 1000 eye-tracker, running at 1000Hz. A nine-point calibration procedure was used, and accepted only when the average error was <0.5 degrees visual angle and no points had an error of >1 degree of visual angle. A drift correct procedure was performed prior to each trial. To limit head movements participants were required to use a chin rest while viewing the monitor from a distance of 71cm in a dimly lit room. Participants gave target-present and target-absent responses using a gamepad response box connected via the USB port. In line with previous studies (e.g. Godwin et al., 2010, 2015) a tone sounded following an incorrect response, allowing participants to construct an awareness of target prevalence.

Stimuli

Participants searched for a T-shape (designated as the target) amongst a set of offset L-shapes. The shapes subtended 1.5˚ of visual angle and were randomly rotated by 0˚, 90˚, 180˚, or 270˚. The shapes were randomly placed on an invisible 5x5 grid, and jittered by a randomly generated distance (up to 2.5˚), in a randomly generated direction within the bounds of their grid cells. On each trial or ‘slide’, the search field contained a total of 16 shapes. Sixteen different colours were used. To ensure salience so that none of the colours visually stood out, they were chosen as approximately equally spaced in CIExyY space, and distributed in a ring in colour space. For the alternating colour condition, colours were chosen in opposite locations in the colour ring to form a pair. This method had been used in previous experiments (e.g. Menneer et al., 2007; Stroud et al., 2012). Equal numbers of distractors of each colour were presented across the experiment, though, on any given slide, the set of distractor colours was randomly selected (note that the distractors could be the same colour as the target). Different participants were given different sets of target colours to search for in order to ensure that the results were not confined to a single set of colours.

Procedure

Testing took place in one session, which lasted approximately 2 hours. Prior to any testing, participants were screened to ensure visual acuity and normal colour vision using Snellen (1862) chart for visual acuity and the Ishihara (1917) colour plates. Participants were encouraged to become aware of the alternating prevalence levels. To do this, pairs of searches were presented to participants as belonging to a single over-arching 'trial', and individual search displays within each 'trial' comprised of two 'slides' (see Figure 2.1). Both slides were actually separate trials in the conventional sense of a search trial. However, for consistency with the experiment as experienced by the participants, the terminology of referring to a "trial" as comprising of two "slides" will be used. Each slide began with the presentation of a gaze-contingent preview of the target (to serve as a reminder of the target colour) for that slide for 500ms. This preview was presented at the centre of the display to ensure that participants began their search in the same location on each trial.

The experiment consisted of 320 trials, broken down into 640 slides, and was preceded by 10 practice trials. A target was presented on 50% of slides, with only a single target appearing in target-present slides. Slides consisted of an alternating sequence of high-prevalence (95%) and low-prevalence (5%) search arrays (the sequence was counterbalanced across participants).

Participants determined the presence or absence of a target in each slide and responded accordingly. In the same-colour condition, participants searched for the same target throughout all displays. In the alternating-colour condition, the high-prevalence target and low-prevalence target were selected to be eight steps in colour space away from each other. Participants were instructed to respond as quickly, but as accurately as they could. The current trial or 'slide' ended once the participant made their response.

Figure 2.1 Illustration of a trial sequence in Experiment 1. To encourage participants to become aware of the alternating prevalence levels, pairs of searches were presented to participants as belonging to a single over-arching 'trial', and individual search displays within each 'trial' comprised of two 'slides'. Target prevalence alternated between high and low for each 'slide' within the trial, which was counter balanced across participants (i.e. either high followed by low, or low followed by high). Participants performed in either an same-colour condition (a) or a alternating-colour condition (b). The actual search arrays were comprised of sixteen items. For the illustration, items have been increased in size.

Search width as a predictor for Accuracy (a) and Total Search Time (b). On the basis of the Fourier analysis, an approximation of participants' modal time before changing direction on each axis was computed, using their dominant frequency component. Dividing participant's overall search time by the modal time before changing direction provided an estimate of the number of changes of direction

The study used a mixed factorial design with three independent variables, Colour Condition (between: same-colour, alternating-colour), Prevalence (within: high, low) and Trial Type (within: present, absent).

Results

A series of 2 (Trial type: Present, Absent) X 2 (Colour Condition: Same-colour, Alternating-colour) X 2 (Prevalence: High, Low) ANOVAs were examined.

Behavioural Analysis

Response Accuracy

In line with previous experiments (e.g. Godwin et al., 2010, 2015) Response accuracy scores were arcsine transformed to normalise the data. The main effect of Trial Type was significant ($F(1, 14) = 13.513$, $p = .002$, ges = .143) with higher accuracy on target absent than target present trials (see Table 2.1). The main effects of Colour Condition and Prevalence were not significant ($F(1, 14) = .655$, $p = .432$, ges = .026; $F(1, 14) = .657$, $p = .431$, ges = .004).

There were significant interactions between Prevalence and Trial Type ($F(1, 14) = 21.111$, $p < .001$, ges = .221) and Prevalence, Trial Type and Colour Condition ($F(1, 14) = 12.295$, $p = .003$, ges = .142). The interactions between Colour Condition and Trial Type ($F(1, 14) = 0$, $p = .999$, ges = 0) or Colour Condition and Prevalence ($F(1, 14) = .014$, $p = .907$, ges = 0) did not reach significance.

Two further 2 (Colour Condition) X 2 (Prevalence) ANOVAs were run to explore the three-way interaction split down by Trial Type. The interactions were significant on both target present trials ($F(1, 14) = 11.719$, $p = .004$, ges = .146) and target absent trials ($F(1, 14) = 7.008$, $p = .019$, ges = .138). In the alternating condition, accuracy was higher in high than low prevalence trials on target present trials and in low than high prevalence trials on target absent trials ($t(7) = 5.328$, $p = .001$; $t(7) = -3.285$, $p = .013$). These comparisons were not significant in the same-colour condition ($t(7) = 1.408$, $p = .202$; $t(7) = -.353$, $p = .734$ respectively).

Reaction Times

There was a significant main effect of Trial Type ($F(1, 14) = 57.879$, $p < .001$, ges = .23) with longer reaction times on target present trials than target absent trials, and Prevalence ($F(1, 14) = 6.087$, $p = .027$, ges = .018), with longer reaction times on high prevalence trials than low prevalence trials (see Table 2.1). The main effect of Colour Condition did not reach significance ($F(1, 14) = .153$, $p = .702$, ges = .009).

The interactions between Prevalence and Trial Type ($F(1, 14) = 13.471$, $p = .003$, ges = .024) and Prevalence, Trial Type and Colour Condition ($F(1, 14) = 9.776$, $p = .007$, ges = 017) reached significance but those between Colour Condition and Trial Type ($F(1, 14) = .77$, $p = .395$, ges = .004) and Colour Condition and Prevalence ($F(1, 14) = 1.393$, $p = .258$, ges = .004) did not.

Two further 2 (Colour Condition) X 2 (Prevalence) ANOVAs were run to break down the three-way interaction with respect to Trial Type. Colour Condition and Prevalence interacted on target absent ($F(1, 14) = 5.078$, $p = .041$, ges = .027) but not target present trials ($F(1, 14) = 2.593$, $p = .13$, ges = .008). For target absent trials, reaction times were faster in the low than high prevalence condition in the alternating-colour condition ($t(7) = 2.906$, $p = .022$), but not in the same colour condition ($t(7) = 1.853$, $p = .106$).

Eye Movement Measures

Incorrect-response trials were first of all removed from the analyses. If a fixation fell within 3 degrees of visual angle to the centre of an object it was treated as having landed on that object. Fixations were removed if they were greater than 1200ms or less than 60ms in duration. Additionally, the first fixation following display onset (prior to the launch of the first saccade), as well as any fixations that coincided with, or followed the button-press response were also removed from the data set.

Mean Fixation Durations

The mean fixation duration was defined as the average length of time for a fixation. There were no significant main effects for Colour Condition, Prevalence or Trial Type ($F(1, 22) = .117$, $p = .736$, ges = .005; $F(1, 22) = 2.5$, $p = .128$, ges = .003; $F(1, 22) = .106$, $p = .747$, ges = 0, see Table 2.1), nor was there a significant interaction between Colour Condition and Trial Type ($F(1, 22) = .448$, $p =.51$, ges = .0), Colour Condition and Prevalence ($F(1, 22) = .526$, $p = .476$, ges = 0), Trial Type and Prevalence ($F(1, 22) = 3.203$, $p = .087$, ges = .003) or Colour Condition, Trial Type and Prevalence ($F(1, 22) = .192$, $p = .665$, ges = 0).

Mean Number of Fixations

The mean number of fixations was defined as the total number of fixations made during a trial. There was a significant main effect for Trial Type ($F(1, 22) = 106.821$, $p < .001$, ges = .466), with a higher number of fixations made on target absent trials than target present trials, and a significant main effect for Prevalence ($F(1, 22) = 6.511$, $p = .018$, ges = .027), with a higher number of fixations made on high prevalence trials than low prevalence trials (see Table 2.1). The main effect of Colour Condition was not significant ($F(1, 22) = 4.002$, $p = .965$, ges = 0).

The interactions between Prevalence and Trial type ($F(1, 22) = 6.534$, $p = .018$, ges = .01) and Prevalence, Trial Type and Colour Condition ($F(1, 22) = 4.419$, $p = .047$, ges = .007) reached significance.

Two further 2 (Colour Condition) X 2 (Prevalence) ANOVAs were run to break down the three-way interaction with respect to Trial Type but neither reached significance on target present ($F(1, 22) = 2.593$, $p = .13$, ges = .008) or target absent ($F(1, 22) = 2.91$, $p = .102$, ges = .016) trials. Analysing the three-way interaction with respect to Colour Condition showed a significant interaction between Trial type and Prevalence in the alternating colour condition ($F(1, 11) = 7.887$, $p = .017$, ges = .045) but not the same colour condition ($F(1, 11) = .165$, $p = .692$, ges = .0). In the alternating colour condition, the target was absent more fixations were made on high than low prevalence trials ($t(11) = 2.345$, $p = .038$; see Table 2.1), but there was no difference when the target was present ($t(11) =.72$, $p = .487$).

Proportion of Items Visited

The proportion of items visited was defined as the proportion of items directly fixated on the display. There was a significant main effect for Trial Type ($F(1, 22) = 119.973$, $p < .001$, ges = .515), with a higher proportion of items visited on target absent trials than target present trials (see Table 2.1). The main effects of Colour Condition and Prevalence failed to reach significance ($F(1, 22) = .069$, $p = .795$, ges

= .002; $F(1, 22) = 3.958$, $p = .059$, ges = .028). The interaction between Prevalence and Trial Type ($F(1, 22) = 6.485$, $p = .018$, ges = .01) was significant. Target prevalence influenced the proportion of items visited on target absent ($t(23) = 2.276$, $p = .033$), but not target present trials($t(23) =.875$, $p = .391$). A higher proportion of items were visited on high than low prevalence trials (see Table 2.1).

The interactions between Colour Condition and Trial Type ($F(1, 22) = .1$, $p = .755$, ges = 0), Colour Condition and Prevalence ($F(1, 22) = 1.657$, $p = .211$, ges = .012) or Colour Condition, Trial Type and Prevalence ($F(1, 22) = 2.858$, $p = .105$, ges = .006) did not reach significance.

Table 2.1 Behavioural data for Experiment 1a. Mean scores for each trial type within each group for Response Accuracy, Mean Reaction Times, Mean Fixation Durations, Mean Number of Fixations and Proportion of Items Visited. Note, parentheses indicate Standard Deviation (SD).

	High				Low			
	Target Present		Target Absent		Target Present		Target Absent	
Measure	Same	Alternating	Same	Alternating	Same	Alternating	Same	Alternating
Response Accuracy	.96 (.03)	.98 (.02)	.98 (03)	.92 (.09)	.94 (.06)	.89 (.06)	.99 (.01)	.99 (.01)
Mean Reaction Times (ms)	1051.881 (459.036)	1025.08 (179.124)	1436.135 (536.312)	1706.871 (477.183)	1016.572 (353.094)	1094.721 (128.602)	1365.728 (580.567)	1337.624 (276.176)
Mean Fixation Durations (ms)	159.979 (16.293)	160.084 (28.381)	159.923 (20.253)	164.180 (28.669)	163.574 (22.496)	167.631 (39.105)	159.428 (17.269)	164.985 (29.286)
Mean Number of Fixations	2.91 (1.073)	2.864 (.398)	5.114 (1.682)	5.576 (1.541)	2.746 (.749)	2.721 (.471)	4.864 (1.79)	4.549 (1.332)
Proportion of Items Visited	.14 (.03)	.14 (.02)	.23 (.06)	.25 (.05)	.14 (.03)	.14 (.02)	.22 (.06)	.21 (.06)

Experiment 1b

In Experiment 1b, the study was repeated by holding target prevalence constant across slides. In holding prevalence constant it allows a check that any interaction of colour condition and prevalence found in Experiment 1a does not occur only when colour changes. Reaction times and response accuracy were examined, in addition to a series of eye movement measures.

Method

Participants

Eight undergraduate participants from the University of Southampton (1 male and 7 females, mean age = 20.5, SD = 2.45 years) with normal or corrected-to-normal colour vision took part in the study for either course credit or payment.

Apparatus

The apparatus for Experiment 1b were the same as Experiment 1a.

Stimuli

The stimuli for Experiment 1b were the same as Experiment 1a.

Procedure

The procedure for Experiment 1b was the same as Experiment 1a, but the design varied in that the study used a mixed factorial design with two (rather than three) independent variables, Colour Condition (between: same-colour, alternating-colour) and Trial Type (within: present, absent).

Results

Medium Prevalence Levels

A series of 2 (Colour Condition: Same-colour, alternating-colour) X 2 (Trial type: Target Present, Target Absent) ANOVAs were examined.

Behavioural Analysis

Response Accuracy

The main effect of Trial Type was significant ($F(1, 6) = 15.608$, $p = .008$, ges = .602), with higher accuracy on Target Absent than Target Present trials (see Table 2.2). The main effect of Colour Condition was not significant ($F(1, 6) = 3.821$, $p = .098$, ges = .21) nor was there a significant interaction between Colour Condition and Trial Type ($F(1, 6) = 1.53$ $p = .262$, ges = .129).

Reaction Times

The main effect of Trial Type was significant ($F(1, 6) = 25.646$, $p = .002$, ges = .384), with longer reaction times on Target Absent trials compared to Target Present trials (see Table 2.2). The main effect of Colour Condition was not significant ($F(1, 6) = .162$, $p = .701$, ges = .023) nor was there a significant interaction between Colour Condition and Trial Type ($F(1, 6) = .303$ $p = .602$, ges = .007).

Eye Movement Measures

Mean Fixation Durations

There were no significant main effects for Colour Condition or Trial Type ($F(1, 6) = .594$, $p = .47$, ges = .085; $F(1, 6) = .323$, $p = .59$, ges = .003), nor was there a significant interaction between Colour Condition and Trial Type ($F(1, 6) = .037$, $p =.853$, ges = 0; see Table 2.2).

Mean Number of Fixations

The main effect of Trial Type was significant ($F(1, 6) = 38.047$, $p < .001$, ges = .612), with a higher number of fixations on Target Absent trials compared to Target Present trials (see Table 2.2). The main effect of Colour Condition was not significant ($F(1, 6) = .043$, $p = .842$, ges = .005) nor was there a significant interaction between Colour Condition and Trial Type ($F(1, 6) = .865$ $p = .388$, ges = .035).

Proportion of Items Visited

The main effect of Trial Type was significant ($F(1, 6) = 38.489$, $p < .001$, ges = .584), with a higher proportion of items visited on Target Absent trials compared to Target Present trials (see Table 2.2). The main effect of Colour Condition was not significant ($F(1, 6) = .002$, $p = .966$, ges = 0) nor was there a significant interaction between Colour Condition and Trial Type ($F(1, 6) = .628$ $p = .458$, ges = .022).

Table 2.2 Behavioural data for Experiment 1b. Mean scores for each trial type for Medium Prevalence for Response Accuracy, Mean Reaction Times, Mean Fixation Durations, Mean number of Fixations and Proportion of Items Visited. Note, parentheses indicate SD.

	Medium			
	Target Present		Target Absent	
Measure	Same	Alternating	Same	Alternating
Response Accuracy	.97 (.02)	.94 (.02)	.99 (.01)	.99 (.00)
Mean Reaction Times (ms)	986.017 (107.658)	959.811 (150.867)	1333.445 (232.319)	1239.155 (328.487)
Mean Fixation Durations (ms)	154.642 (11.888)	146.747 (16.778)	152.741 (8.411)	145.811 (17.196)
Mean Number of Fixations	2.759 (.487)	2.944 (.283)	5.066 (1.253)	4.646 (1.231)
Proportion of Items Visited	.14 (.02)	.15 (.01)	.23 (.05)	.22 (.05)

Discussion

In Experiment 1a, participants searched for a T amongst Ls, in trials which alternated between high (95%) and low (5%) prevalence ‘slides’. The use of alternating-colour or same-colour condition between participants meant that participants could form *expectations* about target prevalence in both colour conditions, but *repetition* benefits would only occur in the same-colour condition, whereby the low prevalence target benefits from the repetition of the high prevalence target (as they are the same colour).

There were three main findings. First, consistent with a wealth of visual search literature, absent trials were responded to differently from present trials (higher accuracy, longer reaction times, a higher proportion of items visited and a higher number of fixations made on target absent trials compared to target present trials). See Chan and Hayward (2013) or Eckstein (2011) for recent reviews on visual search behaviour.

Second, there is clear evidence in support of bottom-up priming influencing the prevalence effect. There was a much stronger effect of prevalence in both behavioural and eye movement measures in the alternating than same-colour condition. This was consistent with Lau and Huang (2010) and Ishibashi et al. (2012), who also found that the prevalence effect could be modulated through repetition of the target.

Third, while unable to determine evidence of top-down expectations influencing visual search, one finding is consistent with it. Only in the proportion of items fixated is there evidence of an influence of prevalence on absent trials, that does not interact with colour condition. While important, the evidence of a top-down influence on prevalence in based on a null effect: that of failing to find an interaction between colour-condition and prevalence. It is always difficult to argue with confidence for the importance of a null result and for this reason, so we consider that Experiment 1a has found very weak evidence of a top-down influence on the prevalence effect.

In Experiment 1b, participants searched for a T amongst Ls, in trials which were fixed at medium (50%) prevalence (i.e. both slides in each trial were the same). No effects were found for the colour condition for any of the behavioural or eye movement measures examined (which were the same as in Experiment 1a). Furthermore, the interaction between prevalence and trial type reported in Experiment 1a disappeared when prevalence was not manipulated. This confirmed the prediction that without differing prevalence levels, there would be no differences between the same-colour and alternating-colour conditions.

Experiment 1b, like Experiment 1a, has shown different responses to absent then present trials that are like those found in much visual search research; higher accuracy, longer reaction times, a higher number of fixations and a higher proportion of items visited on target absent compared to target present trials. In contrast, to Experiment 1a, Experiment 1b found no evidence of differences between Colour Conditions for any of the measures examined. The critical conclusion to be taken from Experiment 1 is that the prevalence effect can be moderated by bottom-up target repetition.

Experiment 2

Experiments 1a and 1b found limited evidence of a role of top-down information in the prevalence effect. In Experiment 2, we took a different approach to examining top-down influences on search, by examining how search is influenced by an individual's expectations of a second searcher's performance. Given the dyadic nature of paired or team searches, expectations surrounding a partner form a core component of what differentiates paired searches from individual searches. Moreover, any shifts in search performance that arise upon the basis of a partner's performance can only be a top-down influence. Put another way, Experiment 2 asked the question: Do top-down expectations play a role in modulating behaviour when the expectation is based on another's behaviour?

As in Experiment 1, participants searched for coloured Ts amongst Ls. However, in Experiment 2, the slides or prevalence of targets were not manipulated but participants were told that a search 'partner' had already searched the display for a second coloured T-shape. Following each trial, feedback was given indicating how both the participant and the (non-existent) 'partner' performed for that particular trial. Although the target prevalence was not directly manipulated to alter expectations in Experiment 2, instead the *perception* of target prevalence was manipulated through manipulating response bias, allowing the participant to build expectations about their 'partner'.

The simulated 'partner's' search performance was manipulated using high, medium or low criterion conditions derived from Signal Detection Theory (Macmillan & Creelman, 2004). In the high criterion condition, the 'partner' had a default tendency to respond 'absent' regardless of whether there was a target present or not.

In the low criterion condition, the ‘partner’ had a default tendency to respond ‘present’, regardless of whether there was a target present or not. The criterion for the medium condition fell between the high and low criterion, and acted as a baseline. Critically, the proportion of actual errors made (averaged across absent and present trials) remained constant across conditions.

It is worth noting that, in the high criterion condition (where the ‘partner’ makes more errors on target present trials, i.e. misses) participants were able to make a correction by detecting the target that the ‘partner’ had missed. In that case, the pair were successful on that trial. In contrast, in the low criterion condition (where the ‘partner’ made more errors on target absent trials, i.e. false alarms) the participant was unable to correct the ‘partner’s’ error, as the detection of a non-present target could not be corrected or reversed. This asymmetry was allowed in order to retain the opportunity for participants to ensure the visual search task was completed as accurately as possible by the pair. From a typical paired search standpoint, this mirrors real-world paired search tasks wherein both members of a pair are punished or rewarded for their search performance.

Experiment 2 explores how participants’ search behaviour is influenced by the nature of their ‘partner’s’ search performance i.e. how does passive viewing of false alarms and misses made by a ‘partner’ influence your expectations and your own visual search for targets. It was predicted that when the ‘partner’ missed targets (high criterion condition) the participant will make more effort to search the display (in terms of longer search times and a higher number of longer fixations) compared to when the ‘partner’ is making false alarms (low criterion condition). Reaction times and response accuracy will be examined, in addition to a series of eye movement measures.

Method

Participants

Thirty-six undergraduate participants (11 males and 25 females, mean age = 23.5 years, SD = 2.95 years) with normal or corrected-to-normal colour vision took part in the Experiment 2 for course credit or payment.

Apparatus

The apparatus was the same as in Experiment 1, using the same monitor, eye-tracker and set-up.

Stimuli

The stimuli were the same as in Experiment 1. As with the pairs of colours selected in the alternating colour condition in Experiment 1, the pairs of colours selected for the participant and their 'partner' were also located on opposite sides of the colour ring.

Procedure

Testing took place over one session, which lasted approximately 2 hours. Prior to any testing, participants in Experiment 2 were also screened to ensure visual acuity and normal colour vision using Snellen (1862) chart for visual acuity and the Ishihara (1917) colour plates. Experiment 2 consisted of 640 trials, preceded by 20 practice trials. A single target was presented on 50% of the trials. Of the 320 target present trials, 160 displayed Target A (the participant's target), and 160 displayed Target B (the 'partner's' target). Each trial began with the presentation of a 500ms gaze-contingent preview of the two targets (both the participants and their 'partner's'), which was presented at the centre of the display to ensure participants started their search in the same location on each trial (see Figure 2.2 for an illustrative trial sequence).

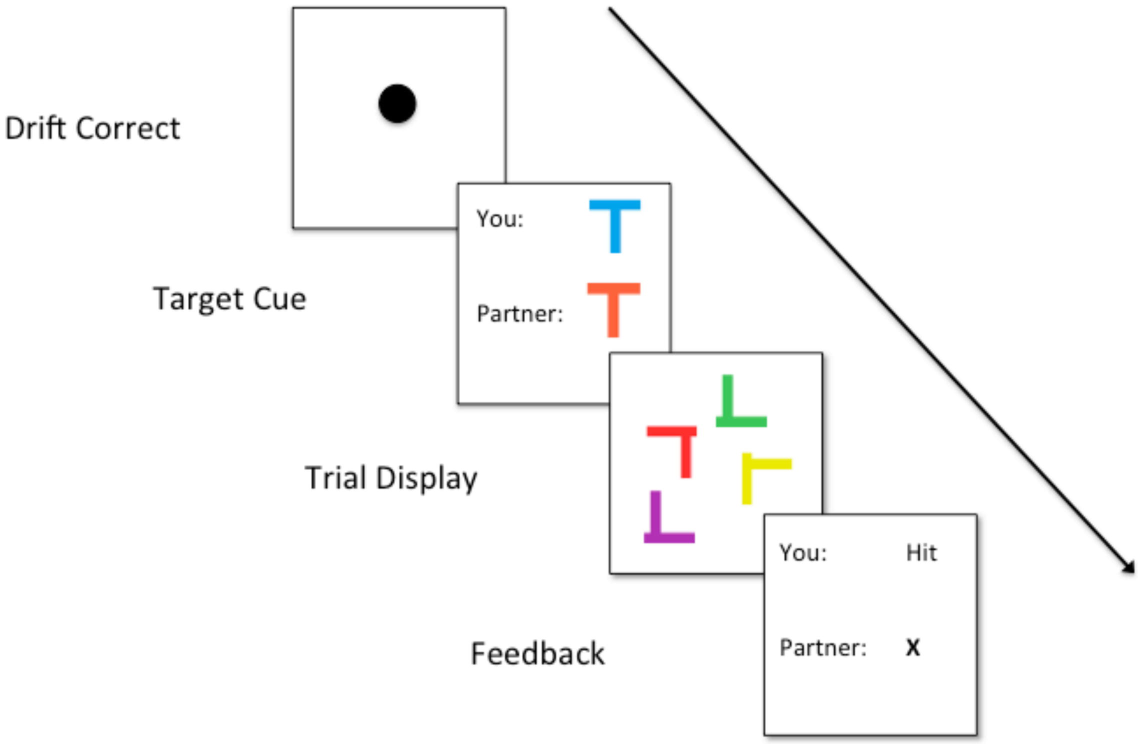

Figure 2.2 Illustration of the trial sequence in Experiment 2. The actual search arrays were comprised of sixteen items. For the illustration, items have been increased in size.

Participants were informed that it was a paired search task, and explained the context in which paired search tasks occur. They were told about the implications of the dual-target cost (Menneer et al., 2007), and that because of the known problems with dual-target search, the study was being carried out with an interest to explore the issues that could arise from splitting up search and combining the search for two different targets from two searchers. They were therefore told that an initial searcher had already completed the task. They were instructed to complete the task with a focus on one of the targets in the pair, as the previous participant had already searched for the other. They were informed, however, that theirs and their 'partner's' results would be combined for analysis of accuracy. The participants were told that their 'partner' would have been allocated one of the two coloured targets to search for, Target B for example, and they were allocated the other target, Target A.

Feedback was given at the end of each trial in the form of a grid, indicating how both the participant and the 'partner' performed on the given trial. During the practice trials, this feedback grid included additional text, and was presented to the participant for 10 seconds (Figure 2.3). This allowed participants the opportunity to familiarise themselves with the feedback system. In the main trials of the experiment, feedback was displayed for two seconds and shown without the additional text. On a Target A trial, the participant received a tick for locating their own target, and a cross for missing it. If they detected their 'partner's' target, they scored a 'hit', but if they did not detect it they were scored a 'miss'. This pattern was reversed for Target B trials. This difference in scoring was included to reinforce that they were to focus on their own target, but could also be scored for their 'partner's'.

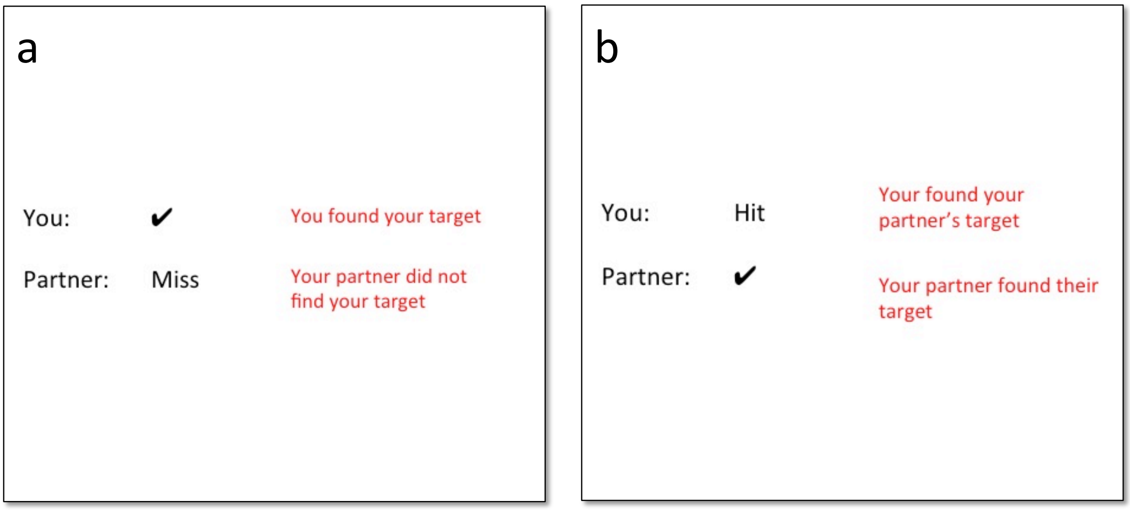

Figure 2.3 Illustrations of a feedback trial. Participants were scored using a tick or a cross for their own target (a), or a hit or a miss for their 'partner's' target (b). Note, the red text was present on practice trials only, to familiarise participants with the feedback system.

The study used a mixed factorial design with two independent variables, Criterion Group (between; high, medium, low criterion) and Trial Type (within; Target A, Target B, Target Absent). The 'partner's' error rate was calculated using response bias measures derived from Signal Detection Theory (Macmillan & Creelman, 2005), with the 'partner's' responses demonstrating either a liberal or conservative bias. In the high criterion condition, the 'partner' would demonstrate a conservative response criterion, e.g. they would respond "absent" frequently. Using a conservative criterion would result in in a higher number of both correct rejections and misses of the target. In the low criterion condition, the 'partner' would shift to hold a liberal response criterion, e.g. they would respond, "present" frequently. Using a liberal criterion would result in a higher number of both target hits and false alarms. The medium condition fell in the middle of the high and low criterion, acting as a baseline. In all three criterion conditions, the sensitivity level remained constant, i.e. the number of errors made remained the same for all conditions.

Similarly to Experiment 1, participants were instructed to respond as quickly, but as accurately as they could. The current trial ended once the participant made their response.

Results

'Partner' Target

For completeness, the number of participants who reported the 'partner's' target (Target B) in addition to their own, are reported. Thirty participants made some attempt to search for Target B, while six participants did not respond to Target B at all. Of these six participants, three were from the medium criterion group, two were from the high criterion group and one was from the low criterion group. However, it cannot be known whether these responses to Target B were responses, or false alarms for Target A. Given the nature of the task, the number of people who responded to Target B was not the focus of Experiment 2, but rather how observing false alarms and misses influences participants' search for their own target. To explore this, a number of 3 (Criterion Group: High Criterion, Medium Criterion, Low Criterion) X 2 (Trial Type: Target Present, Target Absent) ANOVAs were run, with Target Present trials comprised of the 160 Target A trials.

Behavioural Analysis

Response accuracy

The main effect of Trial Type was significant ($F(1, 33) = 196.375$, $p < .001$, ges = .647), with higher accuracy on target absent trials compared to target present trials (see Table 2.3). The main effect of Criterion Group was not significant ($F(2, 33) = 1.382$, $p = .265$, ges = .055) nor was there a significant interaction between Criterion Group and Trial Type ($F(2, 33)$ = .1.382, $p = .326$, ges = .021).

Reaction Times

The main effect of Trial Type was significant (F(1, 33) =86.817, p < .001, ges = .382), with longer reaction times on target absent trials compared to target present trials (see Table 2.3). The main effect of Criterion Group was not significant (F(2, 33) = .429, p = .655, ges = .019) nor was there a significant interaction between Criterion Group and Trial Type (F(2, 33) = .273, p = .763, ges = .004).

Eye Movement Measures

Mean Fixation Duration

The main effect of Trial Type was significant (F(1, 33) = 155.542, p < .001, ges = .243), with longer fixation durations on target absent trials compared to target present trials (see Table 2.3). The main effect of Criterion Group was marginally significant (F(2, 33) = 3.197, p = .054, ges =.153), with longer fixation durations in the high criterion group compared to both the medium (p = .047) and the low (p = .001) criterion group. The interaction between Criterion Group and Trial Type failed to reach significance (F(2, 33) = .816, p =.451, ges =.153).

Mean Number of Fixations

The main effect of Trial Type was significant (F(1, 33) =151.055, p < .001, ges = 553), with more fixations made on target absent trials compared to target present trials (see Table 2.3). The main effect of Criterion Group was not significant (F(2, 33) = 1.321, p = .281, ges = .055) nor was there a significant interaction between Criterion Group and Trial Type (F(2, 33) = .785, p = .464, ges = .013).

Proportion of Items Visited

The main effect of Trial Type was significant (F(1, 33) =182.23, p < .001, ges = 571), with more items visited on target absent trials compared to target present trials (see Table 2.3). The main effect of Criterion Group was not significant (F(2, 33) = .986, p = .384, ges = .043) nor was there a significant interaction between Criterion Group and Trial Type (F(2, 33) = .467, p = .631, ges = .007).

Table 2.3 Behavioural data for Experiment 2. Mean scores for each trial type within each group for Response Accuracy, Mean Reaction Times, Mean Fixation Durations, Mean Number of Fixations and Proportion of Items Visited. Note, parentheses indicate SD.

	Low Criterion		Medium Criterion		High Criterion	
Measures	Target Present	Target Absent	Target Present	Target Absent	Target Present	Target Absent
Response Accuracy	.92 (.05)	.99 (.01)	.91 (.023)	.99 (.004)	.90 (.06)	.99 (.01)
Mean Reaction Times (ms)	1345.607 (276.6)	2469.344 (823.325)	1196.689 208.589	2192.577 (519.7)	1163.280 (244.266)	1920.567 (653.07)
Mean Fixation Durations (ms)	142.931 (14.706)	159.299 (13.538)	151.349 (16.596)	168.641 (17.161)	155.948 (16.774)	173.033 (16.014)
Mean Number of Fixations	3.723 (1.091)	10.297 (.3.386)	3.096 (.734)	8.546 (1.928)	2.753 (.907)	7.180 (2.859)
Proportion of Items Visited	.20 (.05)	.49 (.14)	.17 (.04)	.43 (.09)	.15 (.05)	.37 (.13)

Discussion

Experiment 2 explored top-down expectations in a paired search scenario, examining the influence of searching with an assumed second visual searcher, who searched using either a liberal or conservative response criterion. As in Experiment 1, participants searched for coloured T-Shapes amongst L-Shaped distractors. However, in Experiment 2 prevalence of targets was not manipulated to alter expectations, but instead participants were told that a search 'partner' had already searched the display for a second coloured T-shape. Experiment 2 considered how the participant expected the 'partner' to perform influenced their own search behaviour.

The prediction was that visual search would be influenced by the uncertainties created by observing misses and false alarms. Specifically, it was predicted that observing misses (high criterion condition) would make participants more conservative in their search. Conservative searches leading to longer search times and a higher number of longer fixations compared to when the 'partner' is making false alarms (low criterion condition).

The data show evidence of the difference between absent and present responses that is routinely found in visual search experiments; accuracy was higher, reaction times were longer, and participants made a higher number of longer fixations, and to more items on target absent trials compared to target present trials. In contrast, the behavioural data showed no evidence of an effect of observing the performance of a simulated second searcher on reaction times or accuracy. But that is not to say that visual search behaviour was unaffected by the simulated second searcher. While the evidence is only marginally significant, fixation durations increased as response criterion increased. This effect is that viewing misses led to increased fixation durations. However, there was no evidence to suggest longer search time or an increased number of fixations, therefore it would not be appropriate to conclude that observing misses in a 'partner' leads to more conservative search behaviour. The most accurate conclusion is that observing misses in a search 'partner' creates uncertainty for participants and this changes how search is conducted but not its outcome.

Of course, any conclusions reached here are limited by the strength of the data. The data supporting this conclusion are weak but strong enough to provide reason to pursue this issue further in future work. Most obviously, the simulated searcher used in this experiment may be a rather weak instantiation of a second searcher. Increasing the credibility of the second searcher may provide stronger data in support of the hypothesis tested here.

General Discussion

The dual-target cost suggests that when two types of target are searched for simultaneously, a cost in performance will occur (Menneer et al., 2007). However, if suggestions to split search so that different searchers search for specific target categories within a multiple target type task are to be implemented, the implications of more than one searcher working together must be addressed. Especially given that many real world visual search tasks which require paired searching or teamwork, are safety critical and/or require high levels of efficiency. Furthermore, airports in the UK are already advancing and introducing “Centralised Image Projection” whereby a second searcher checks the performance of an initial searcher, so the implications of Chapter 2 is important in regards to that too. When searching with a partner, expectancies can occur surrounding their performance. An important question, therefore, is do these expectancies influence your own search behaviour.

To establish a baseline for understanding expectancies in individual search, Experiment 1 explored the expectancies that develop when participants search for targets at different prevalence levels. Alternating versus same-colour conditions were created to allow a condition in which bottom-up priming occurred. If the target colour remained the same across high and low prevalence slides, then it was predicted that the prevalence effect would be alleviated due to the influence of bottom-up priming. Experiment 1’s prediction was supported with a stronger prevalence effect found in the alternating-than same-colour condition, in both behavioural and eye movement data. The finding demonstrates the importance of target identity remaining constant if the prevalence effect is to be mitigated.

Experiment 2 explored the expectancies that can develop when participants search with a ‘partner’, who demonstrates a response criterion. It is important to note, that the focus of Experiment 2 was how witnessing someone else’s search behaviour influences your own. It was predicted that when the ‘partner’ made more misses (high criterion condition) the participant would make more effort to search the display in terms of longer search times and a higher number of longer fixations) compared to when the 'partner' is making false alarms (low criterion condition). Experiment 2 revealed that fixation durations increased as response criterion increased.

The implications of the findings from Experiments 1 and 2 are important. Dealing with targets that might appear with two identities leads to an enhanced sensitivity to the prevalence effect. This finding is rather similar to the weaker guidance that results from having to search for two target types simultaneously (Menneer et al., 2007). Attending to the second target may also interrupt the bottom-up priming that comes from the repetition of your own target, as demonstrated in the results of Experiment 1.

The fact that the significant result was found when the 'partner' was making misses also suggests that the competence of the additional searcher can have an influence on search strategy. Others have started to explore the implications of working together in terms of efficiency as a pair (i.e. when you are required to coordinate, not search for two target types separately). Previous studies have found pairs to be quicker (and as accurate) at detecting and identifying targets than participants performing the task alone (Brennan et al., 2008; Wolfe et al., 2007), but have pointed out that collaborative search incurs coordination costs in terms of who searches where, and when to end target absent trials (Brennan et al., 2008). Furthermore, Wolfe et al. (2007) noted that it may be hard to distinguish whether performance in a paired task is due to the collaborate effort of the pair, or whether it reflects the performance of the superior participant.

When a task requires a searcher to focus on a single target, but requires you to work simultaneously with another searcher, the present experiments suggest it is important to encourage participants to actively ignore other targets and the quality of search being conducted by a partner searcher.

Chapter 3

Searching for Targets in Grassland: The Importance of Search Strategy and Working Independently

In terms of the complex search task of searching for clues to a crime or IEDs in patrol on a large area of grassland, Chapter 1 revealed a number of factors to be important. In addition to the issues explored in Chapter 2, this form of search task is also large-scale, requiring eye, head and body movement (Foulsham et al., 2014; Gilchrist et al., 2001; Jiang et al., 2014), integrates aspects of foraging and searching for an unknown number of targets (Cain et al., 2012; Wolfe, 2013) and requires defining and following a search path (Koopman, 1946, 1980). Furthermore, the implication of paired searching is also of importance. Chapter 3 took these factors and explored them in relation to how working in a pair might influence search in a task comparable to the police searching for clues to a crime, or soldiers searching for IEDs in patrol.

What do we know about how police search teams comb open ground for clues to a crime, or how soldiers patrolling high-risk routes search for Improvised Explosive Devices (IEDs)? The unfortunate answer is that, at present, very little is understood. With a few exceptions (e.g. Foulsham et al., 2014; Gilchrist et al., 2001; Jiang et al., 2014), previous studies on human search ability have primarily explored visual search (see Eckstein, 2011 for a recent review) or foraging (Cain et al., 2012; Wolfe, 2013) using experimental stimuli presented on computer screens.

There are three major differences that should be noted between computer-based tasks and the task of searching for small targets placed in open space. First, the spatial scale is sufficiently different that small targets in open spaces are unlikely to be detected through pre-attentive or attentive vision alone (e.g. Treisman & Gelade, 1980; Wolfe, 2007; Wolfe, Cave, & Franzel, 1989) but will also require head and body movement (i.e., physical foraging behaviour). Second, unlike terminating search following the detection of a single target (Tuddenham, 1962a), or when the rewards of continuing foraging in one area are less than those that might follow from moving to a new area (Cain et al., 2012; Wolfe, 2013), searchers must attempt to search spatial locations exhaustively (i.e. try their best to search as many possible target locations as they can). Third, if you consider real world situations such as police searching for clues to a crime or soldiers searching for IEDs, searchers have to consider all potential instances of all potential target types (Godwin et al., 2015), rather than finding well-defined targets matching a simple template. Chapter 3 is a first effort in trying to understand what leads to good search performance in these types of task. The focus is limited to understanding the search strategies that maximize the chances of finding all instances of an unknown number of targets, which are distributed, across an area of open space.

Search strategies have previously been explored in studies of eye movements, in terms of fixation patterns in active vision tasks (e.g. Gilchrist & Harvey, 2006; Keech & Resca, 2010). They have also been explored in terms of the approach undertaken by typical individuals and brain-damaged patients when performing computerized versions of pencil and paper cancellation tasks (Dalmaijer, Van der Stigchel, Nijboer, Cornelissen, & Husain, 2015; Donnelly et al., 1999). Despite differences in tasks and goals across these studies, a common conclusion is that detecting targets is better when search follows specific, structured paths.

The utility of using systematic paths for search is also apparent in mathematical models aimed at optimizing real-world search. Search Theory (Koopman, 1946, 1980) was developed in the Second World War to facilitate maritime search, rescue and detection operations. The theory is an application of probability theory to search such that the likelihood of targets being found at specific locations can be computed, enabling searching of areas likely to contain targets to be prioritized. On the basis of these calculations, a search path is defined over which a plane or ship can pass in search of targets, with an improved chance of detecting them. Search theory is still being developed, with recent applications for search and rescue on land (Koester, Chiacchia, & Twardy, 2014; Robe & Frost, 2002). Search theory can be utilized for simple through to complex search scenarios. In the simplest of cases, the likelihood of targets appearing within a search area might follow a uniform distribution. More complex cases must take account of Gaussian distribution of probabilities for factors such as initial target locations, drift patterns caused by air and water currents, and perhaps a desire for targets to remain hidden.

Considering the most simple of cases, optimal paths should minimize the distance of the route that allows all locations to be searched (using body, head and eye movements) without making revisits. Evidence can be found in support of humans using such a strategy in Gilchrist et al., (2001). They explored whether the principles from visual search studies could be applied to large-scale search tasks. Using a task where all potential target locations were set in a regular array and clearly visible on the ground, they examined how the search for marble targets hidden in film canisters was conducted. The important result was that, in contrast to visual search, rechecking locations was rare with participants making fewer revisits to canisters they had already searched. Gilchrist et al. (2001) suggested that there was a higher cost (in terms of the physical effort to cross the room) associated with forgetting in their task relative to standard visual search, and that the increased effort required to search led to a higher likelihood of participants remembering searched locations. A similar result was reported by Smith, Hood and Gilchrist (2008) in the same large-scale experimental setting, and also by Ruddle and Lessels (2006). In all these tasks, minimizing re-checking must have involved determining a route through the search array.

Gilchrist et al. (2001) used a regular, visible search array (although targets were hidden in film canisters and required checking, these canisters were clearly laid out on the floor), but what happens when targets are hard to find and potential target locations are not arranged in a regular fashion? Critical to understanding how humans might conduct effective target search in such circumstances is what is meant by trying to search exhaustively across space. Consider the case where a given search area is wider than that which can be searched in a single pass. Within Search Theory (Koopman, 1943; 1980), the distance to the left and right of each location that still allows target detection defines an area known as the effective search width (ESW). The technical limit of sensors (human vision, radar etc.) determines the width of the ESW and therefore how close together neighbouring passes of a single path should be. With respect to humans, the width of the search corridor is referred to rather than the ESW, as human search is not deterministic. Whether a particular setting of search width is considered effective for target search is dependent on knowing the accuracy of target detection within a search corridor.

For humans, without making head and eye movements to the left and right, the search width will be determined by the limits of foveal and parafoveal vision to discriminate targets. However, head and eye movements can be made to overcome the limits of foveal vision so long as they are calibrated with the forward speed of travel (along the search corridor being 'swept'). If the speed of sweeping along a search path is too fast to allow search of adjoining spaces to the left, right and centre then search will be incomplete. In conclusion, the likelihood of search being conducted in a fine-grained and exhaustive manner will depend on a range of factors. In the absence of existing data, it seems likely that search will be subject to the strategic decisions made by searchers regarding systematicity of search, and the trade-off of their chosen width of search corridor for sweeping, with their forward speed of search. Experiment 1a was designed to reveal evidence of individuals using a search strategy (in terms of the forward speed of search, the width of the corridor being swept and consistency of a chosen strategy) and to assess the influence of this strategy on search accuracy and overall speed.

Experiment 1a

In Experiment 1a, individual participants searched for an unknown number of coin targets placed in open grassland terrain. The grassland was 75m^2 in size. Participants searched the open space in any manner they chose, and for however long they wished to search. In Chapter 3, the interest is in how search strategy was related to both target detection accuracy and search time for the task. Accuracy was measured as the number of targets detected, search time was calculated as the time from when the participants started the task until they told the experimenters they had finished.

In addition, analysis of the systematicity of search was enhanced by using data extracted from a Total Station theodolite system. These data allow for visual representation of the routes taken by participants (henceforth route maps: see

Figure 3.3). They can also be processed using Fourier analysis to provide some quantitative evidence for general trends apparent when inspecting representations of routes taken by participants. This approach is helpful as participants are free to move in any direction and it captures the underlying spatio-temporal properties of their movements overall.

The Fourier analysis transforms the data from the time domain to the frequency domain and enables calculation of multiple measures derived from the frequency components of movement along the X- and Y-axes of the search area. (1) The dominant frequency component is an index of the modal speed of movement along each axis. The reciprocal of the dominant frequency component (1/dominant frequency) converts this to seconds, and can be used as an approximate measure of participant's modal time before changing direction (2) Dividing participant's overall search time by the modal time before changing direction gives an estimate of the number of changes of direction. Plotting the number of changes of direction along the X-axis against those on the Y-axis allows determination of whether participants tended to move systematically along X- or Y-axes (i.e., left-to-right or top-to-bottom), or use a hybrid strategy. Furthermore, by dividing the number of changes of direction by the length of the axis being travelled across provides an estimate of the width of the search corridor used by participants. Finally, (3) dividing participants' modal time before changing direction by the next most commonly occurring time before changing direction indexes variability in speed of searching.

In addition to basic data around search accuracy and time, typical route maps are shown, and the width of search corridors used, modal movement speed and variability in speed of movement are analysed. It was predicted that more accurate search would be reliant on (1) increased regularity, systematicity of search following a structured path (Dalmaijer et al., 2015; Donnelly et al., 1999; Gilchrist & Harvey, 2006; Keech & Resca, 2010), (2) narrower search corridors (Koopman, 1943, 1980) and (3) slower, more consistently paced movement (Koopman, 1943, 1980).

Method

Participants

Thirty participants (7 males and 23 females, mean age=23.5 years, SD=4.73) recruited from the University of Southampton community, with normal or corrected-to-normal colour vision, took part in the study for course credit. Participants were screened to ensure visual acuity and normal colour vision using Snellen (1862) chart for visual acuity and the Ishihara (1917) colour plates. The study was performed in accordance with the Declaration of Helsinki and was approved by the University of

Southampton, School of Psychology ethics committee; informed consent was obtained from all participants.

Apparatus

Experimenters used a grid representing the larger scale search space to record accuracy. When participants found a target they pointed to the target and informed the experimenter, who then marked off the corresponding target on the grid. Overall search time was recorded using a stopwatch. Search was deemed to have finished once participants reported to the experimenter that they were done.

Participant movement over space and time was recorded using a Total Station theodolite (Leica TPS1200, Heerbrugg, Switzerland). The Total Station used electronic distance measurement technology and an angle-measurement system to calculate the coordinate of an unknown point relative to a known coordinate point. A signal was sent from a fixed recording station to a reflector prism mounted on a 1.8 meter staff held by the participants. A coordinate was recorded every two seconds and accuracy was within 3 mm per kilometre of distance (SD=1.5). The output of time stamped coordinates was processed using Environmental Systems Research Institute's ArcGIS software (ESRI, 2011).

Stimuli

The experiment was conducted on an open space of grassland (see Figure 3.1) The perimeter of a 15m x 5m search area was marked out at 1-meter intervals using 40 coloured cones. The position of the cones was calibrated with the Total Station prior to testing. The relative positions of cones allowed definition of 75m^2 grid cells. Testing took place over four consecutive days. The location of the grid was moved each day to avoid excessive trampling of grass. Across days, the conditions of the grass remained broadly similar.

Figure 3.1 The search grid. An example photograph of the 75m² grid in which participants searched for coin targets. The perimeter of a 15m x 5m search area was marked out at 1-meter intervals using 40 coloured cones.

Twenty-five UK sterling two pence coins were used as targets (see Figure 3.2). The two pence coins were of a copper colour and were 25mm in diameter. The coins were all matt rather than shiny in appearance. The coins were placed so that they could be detected from standing height, although they were sufficiently small that detection required active exploration of the search grid.

Coins were distributed within the grid so that 5 coins were placed pseudo randomly within each 1-meter 'search lane' (i.e. 5 coins were placed along each 1-meter search lane across the 5 meter width of the grid). On average there was one coin per $3m^2$. The targets were placed in the same grid cell locations for all participants. Distractors were not added within the search area but leaves and other natural materials did form naturally occurring distractors and were not removed from the search space.

Figure 3.2 A target. An example photograph of a UK sterling two pence coin target set within the search grid. Note the leaves which occurred naturally and acted as distractors.

Procedure

Following the screening, participants were given a brief explanation of the Total Station and were instructed to hold the staff upright and close to their body while searching. The staff was lightweight and easy to carry. Participants were not told how many coins could be found but instructed to search until they were confident they had completed their search (i.e. they had found all the targets). Participants were told not to pick coins up but to point and tell the experimenter that a coin had been found. Participants were not penalized for reporting the same coin on more than one occasion (as the task simulates a task where a conservative approach to finding targets is encouraged) but each target was only counted once when calculating response accuracy. Once participants had completed their search, they were asked to give a score on a ten-point scale to indicate how confident they were that they had found all the coins. A higher score implied high confidence while a lower score indicated lower levels of confidence.

Results

Participants were excluded from data analysis if they failed to detect any coins. While detecting no coins might reflect their best performance, they were removed on the basis that it is impossible to differentiate poor performance from failing to engage with the task. This resulted in the removal of two participants (6.7% of the data) meaning data analysis was conducted on the data from 28 participants. Correlational tests were used to examine if there was a relationship between two variables, simple linear regressions were used to examine whether one variable predicted a second variable, t-tests were used to examine whether the mean scores of two participant groups differed and a Fisher's exact test was used to test how likely it was that observed distributions were due to chance. All regressions reporting time or frequency use log-transformed data to reduce skew, though this did not, however, affect the underlying pattern of results. For the regressions, significance levels were adjusted for multiple comparisons as regressions compared the same measure across X- and Y-axes. Only effects reaching a p value of .025 were considered significant. The statistical package used to analyse the data was R version 3.3.0 (R Core Team, 2016).

Behavioural Data

Basic measures of accuracy and total search time are presented in Table 3.1. On average, participants found just under half the available targets, despite spending an average of seven and a half minutes on the task. On average, participants reported a confidence score of 6.89 on a ten-point scale. On average, the first target was found after 39 seconds and the last target was found 40 seconds before terminating search. For each participant, a regression was carried out exploring the linear relationship between the time of finding each target against the ordinal number of that target as found by the participant (i.e. 1st, 2nd, 3rd etc.). This measure explores whether targets were found consistently throughout search or were found more easily at the beginning than the end of search. The range of adjusted R-squared values was .842 to .995 (all ps < .052) across participants. The result suggests that targets were found at a fixed rate throughout search. Accuracy was predicted by confidence ratings (β_1 = 2.448, $F(1,27) = 5.195$, $p = .031$, adj $R^2 = .134$): participants who gave a high confidence rating were more accurate in their search.

Table 3.1 Behavioural data for Experiment 1a. The Mean, Minimum and Maximum Scores for Accuracy and Total Search Time for the Individuals in Experiment 1a, the Individuals in the Three Target Frequency Conditions of Experiment 1b and the Dyads in Experiment 2. Note, parentheses indicate SD.

	Accuracy					Total Search Time (min:s)				
	Experiment 1a	Experiment 1b			Experiment 2	Experiment 1a	Experiment 1b			Experiment 2
	Individuals	5 Targets	15 Targets	25 Targets	Dyads	Individuals	5 Targets	15 Targets	25 Targets	Dyads
Mean	0.45 (0.04)	0.76 (0.19)	0.84 (0.16)	0.8 (0.11)	0.68 (0.13)	07:33 (02:32)	03:51 (02:49)	03:41 (02:00)	03:41 (01:40)	06:15 (04:13)
Minimum	0.12	0.4	0.47	0.52	0.4	02:32	01:19	01:38	01:05	01:13
Maximum	0.8	1	1	0.96	0.92	25:08:00	07:35	08:32	11:53	18:05

Search Strategy

Examples of typical route maps are shown in

Figure 3.3. Visual inspection of these route maps reveals some commonalities across participants, i.e. participants tended to search in multiple search lanes (or an 'S' shaped pattern) either top-to-bottom or left-to-right. These commonalties were explored using Fourier analysis.

Systematicity of Search Path

To quantify regularity of search path a Fourier transformation was carried out for each participant and along both the X- and Y- axes (see Donnelly et al., 1999). The first prediction was that search would be more accurate when participants used a regular, systematic search strategy following a structured path. To explore this prediction the number of changes of direction (calculated by dividing participant's overall search time by the modal time before changing direction) were examined. *The number of changes of direction was plotted across both axes (see Figure 3.4). High values on one axis were associated with low values on the other axis ($r(26) = -.558$, $p = .002$). Participants turning top-to-bottom tended to search along fewer, longer corridors and participants turning left-to-right tending to search along more, shorter corridors. These data are consistent with all participants searching systematically, using a 'S' shaped route to cover the search area (as shown in*

Figure 3.3). Their fundamental pattern of movement (the 'S' shape) was consistent irrespective of whether they primarily moved top-to-bottom or left-to-right. The important result is that all participants exhibited regularity in the path taken to search. Given this, it is not possible to explore variations in accuracy as a function of the presence or absence of regularity (as per the first prediction).

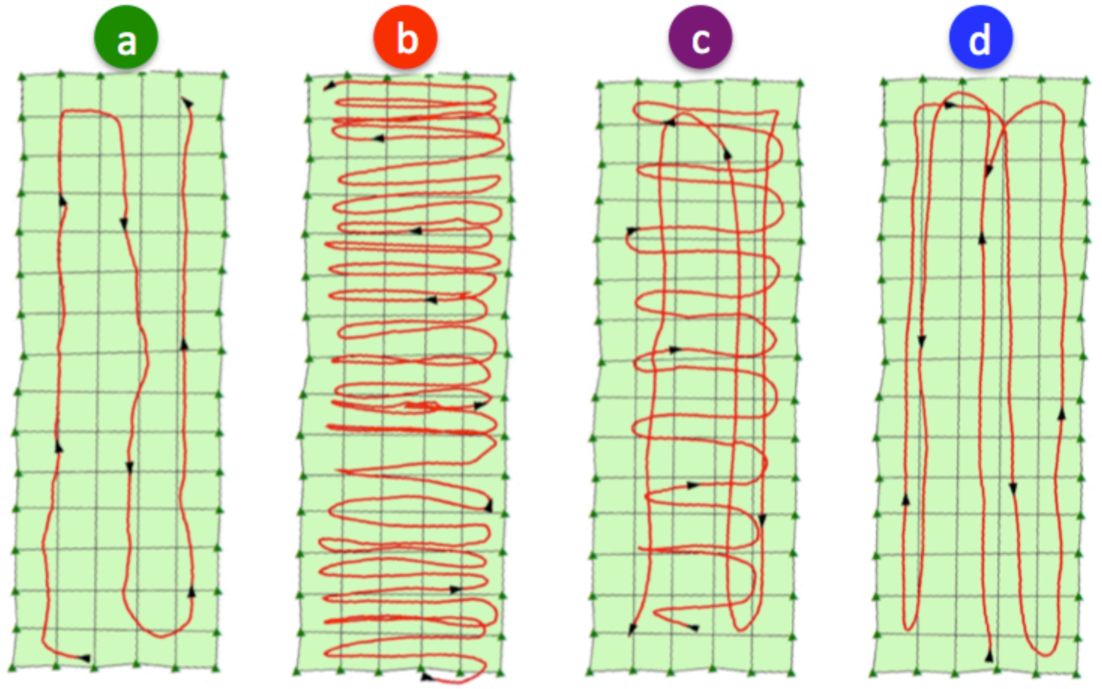

Figure 3.3 Route maps for Experiment 1a. Illustrations of four examples of routes taken when searching. The examples have been chosen to illustrate how the commonly used 'S' shape strategy could occur alone or as part of a more complex search, and fixed along the X- or Y- axes (i.e., searching left-to-right or top-to-bottom). Note, the colours of the labels correspond to the coloured plot points in Figure 3.4.

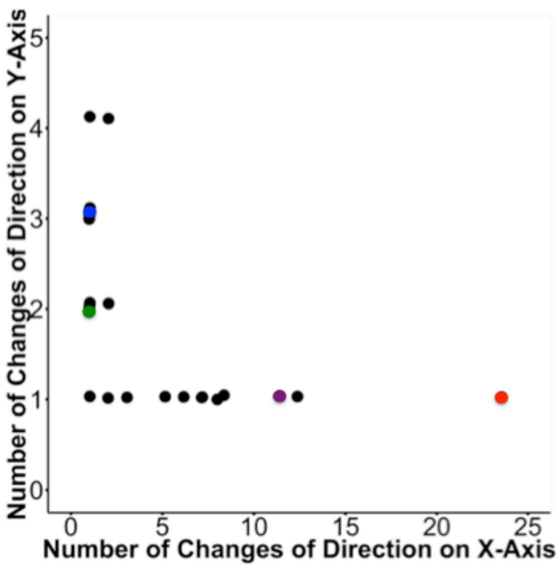

Figure 3.4 The number of changes of direction on each axis. On the basis of the Fourier analysis, participants' modal time before changing direction on each axis was computed, using their dominant frequency component. Dividing participant's overall search time by the modal time before changing direction provided an estimate of the number of changes of direction. Figure 3.4 shows the relationship between the number of changes of direction on the X- and Y- axes. Note, non-logged data is plotted throughout. Note, the coloured plot points correspond to the coloured labels in Figure 3.3.

Width of Search Corridors

The second prediction was that search would be more accurate as search width narrowed. This hypothesis was based on the idea that a narrow search width would better facilitate the search of close space using the fine-grained spatial acuity of foveal vision. To explore this prediction participants were split into two groups – those who primarily moved left-to-right (n = 11) and those who primarily moved top-to-bottom (n = 16). There was one participant who did not fall into either category, conducting a hybrid strategy, and was therefore removed from subsequent analysis. The number of changes of direction made by each participant along their dominant axis (i.e., whether they were in the top-to-bottom or left-to-right group), was scored, added 1 to (to take into account the number of sweeps both up and down, i.e., 5 turns would mean 6 sweeps) and then the length of the axis being travelled across was divided by this figure (i.e., for those in the top-to-bottom group, the length of the axis being travelled across in meters, which was 5, would be divided by the first figure calculated). This normalized the data, as calculating the search width took into account the length of each axis. These data are shown in Figure 3.5.

The striking result is that the search width for the majority of participants lies between 1 and 2 meters (alternatively between 50 cm and 1m to both the left and right of the centre). This suggests that there is commonality in search width irrespective of whether participants search top to bottom or left to right across the search grid. Given the limited range of width of search corridors, there is no evidence of search width predicting either accuracy ($\beta_1 = -.534$, $F(1,25) = .356$, $p = .556$, adj $R^2 = -.025$), or total search time ($\beta_1 = .446$ $F(1,25) = 2.11$, $p = .159$, adj $R^2 = .041$). Irrespective of outcome for accuracy or time, participants searched along their ‘S’ shaped path, using a common search width.

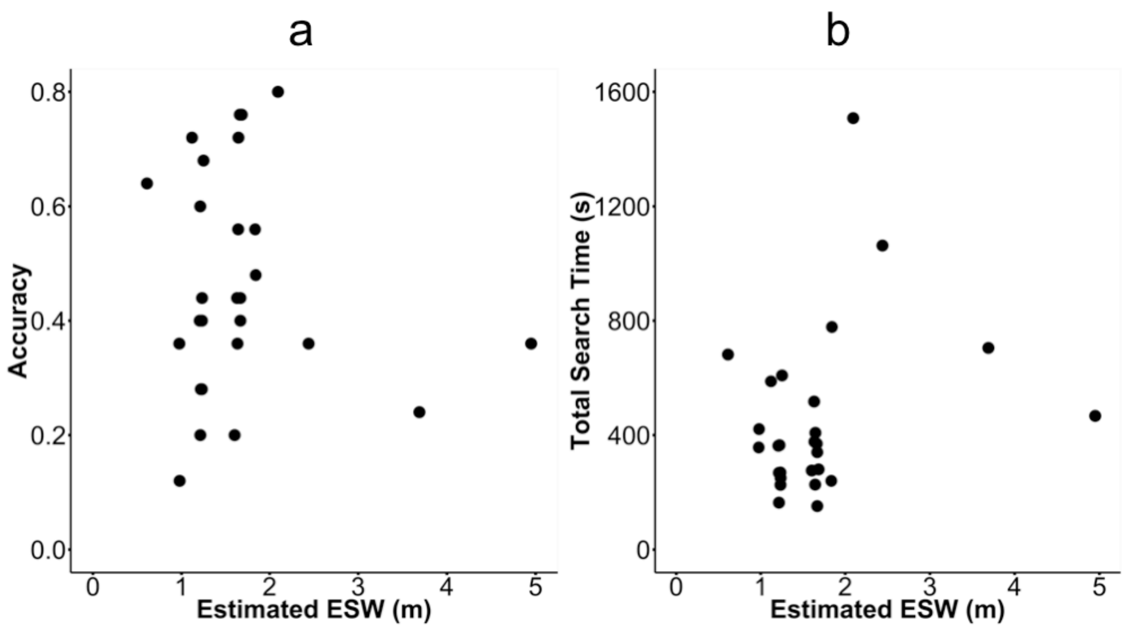

Figure 3.5 Search width as a predictor for Accuracy (a) and Total Search Time (b). On the basis of the Fourier analysis, an approximation of participants' modal time before changing direction on each axis was computed, using their dominant frequency component. Dividing participant's overall search time by the modal time before changing direction provided an estimate of the number of changes of direction To take into account the number of sweeps both up and down, 1 was added to the number of changes of direction made by each participant along their dominant axis (i.e., whether they were in the left-to-right, or top-to-bottom search group) and the length of the axis being travelled across was divided by this figure. This provided a normalized search width in meters (m). Accuracy was the proportion of coins detected and Total Search Time indicates the total amount of time spent searching in seconds (s).

Search Speed and Search Speed Variability

The third prediction was that search would be more accurate when participants searched with slow, consistently paced movement (Search Theory; Koopman, 1943; 1980). To examine this prediction the modal time before changing direction on each axes, and variability in time before changing direction (by dividing the modal time before changing direction by the next most commonly occurring time before changing direction) was examined. For the left-to-right group, accuracy was not predicted by the modal time before changing direction, nor variability in time before changing direction (β_1 = -.41, $F(1,9)$ = .065, p =.805, adj R^2 = -.103; β_1 = -3.437, $F(1,9)$ = -1.478, p =.255, adj R^2 = .056; see Figure 3.6a-b). Total search time was predicted by modal time before changing direction but not variability in time before changing direction (β_1 = -1.133, $F(1,9)$ = 8.714, p =.016, adj R^2 = .436; β_1 = .094, $F(1,9)$ = .009, p =.928, adj R^2 = -.11; see Figure 3.6c-d).

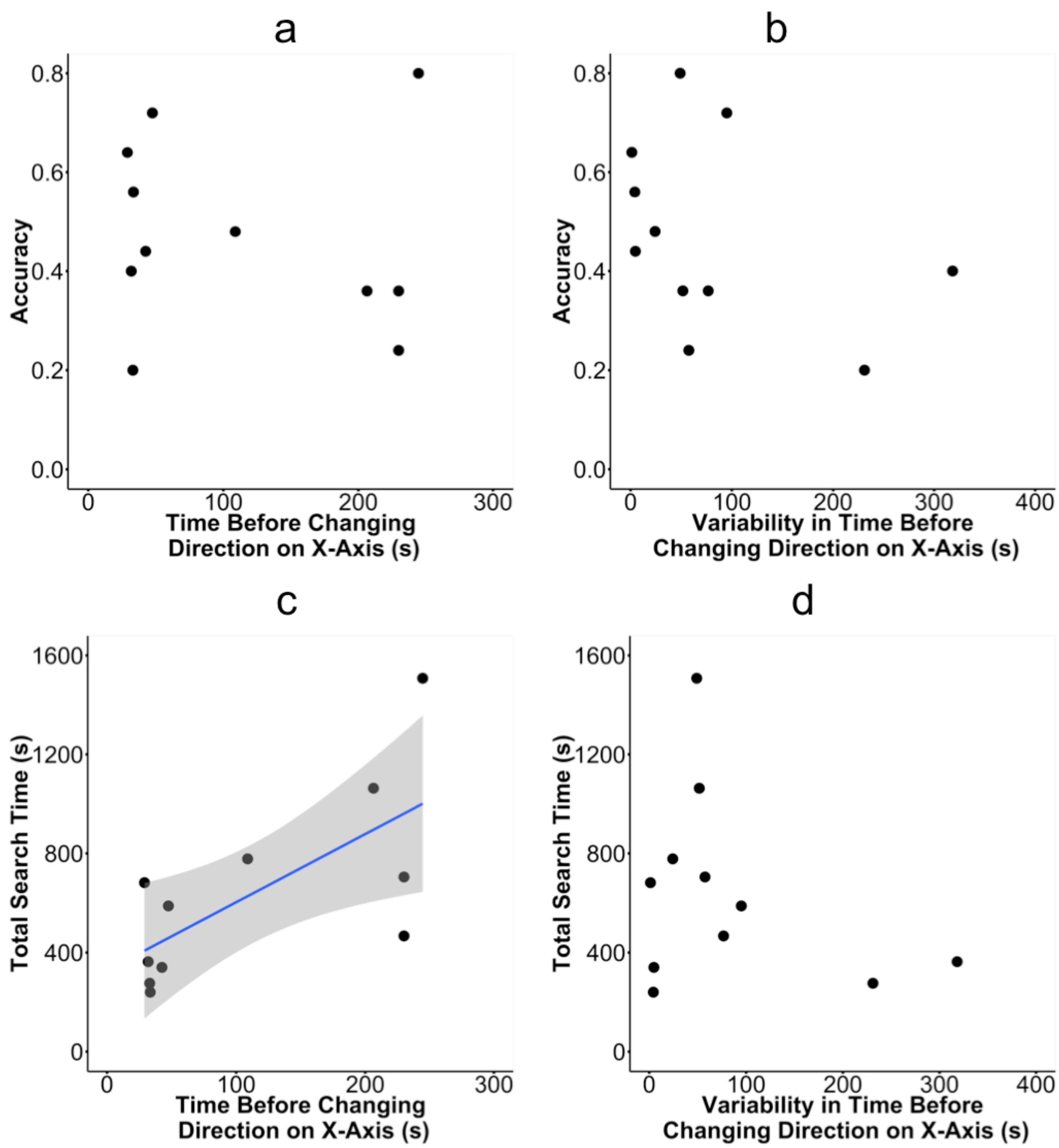

Figure 3.6 Search speed to predict Accuracy and Total Search Time for left-to-right participants on the X-axis. On the basis of the Fourier analysis, an approximation of participants' modal time before changing direction in seconds (s) on each axis was computed, using their dominant frequency component. By dividing the modal time before changing direction by the next most commonly occurring time before changing direction, the variability in time before changing direction in seconds (s) was computed. Accuracy was the proportion of coins detected and Total Search Time indicates the total amount of time spent searching in seconds (s). Figure 3.6 shows the modal time before changing direction (a) and the variability in time before changing direction (b) to predict Accuracy. Plus the modal time before changing direction (c) and the variability in time before changing direction (d) to predict Total Search time.

For the top-to-bottom group, there was a very strong trend for accuracy to be predicted by modal time before changing direction, and variability in time before changing direction (β_1 = 1.229, $F(1,14)$ = 5.926, p = .029, adj R^2 = .247; β_1 = 1.982, $F(1,14)$ = 5.716, p =.031, adj R^2 = .239; see Figure 3.7a-b). Participants were more accurate when they took longer before changing direction, and varied their time before changing direction. Total search time was predicted by modal time before changing direction but not variability in time before changing direction (β_1 = -0.956, $F(1,14)$ = 27.21, p < .001, adj R^2 = .636; β_1 = .507, $F(1,14)$ = 1.047, p = .324, adj R^2 = .003; see Figure 3.7c-d). Participants took longer overall to search when they took longer before changing direction. As predicted, increased accuracy was associated with slow search (i.e., longer before changing direction, i.e., turning), but surprisingly, it was variable search speed that was associated with increased accuracy, rather than a consistent pace as predicted.

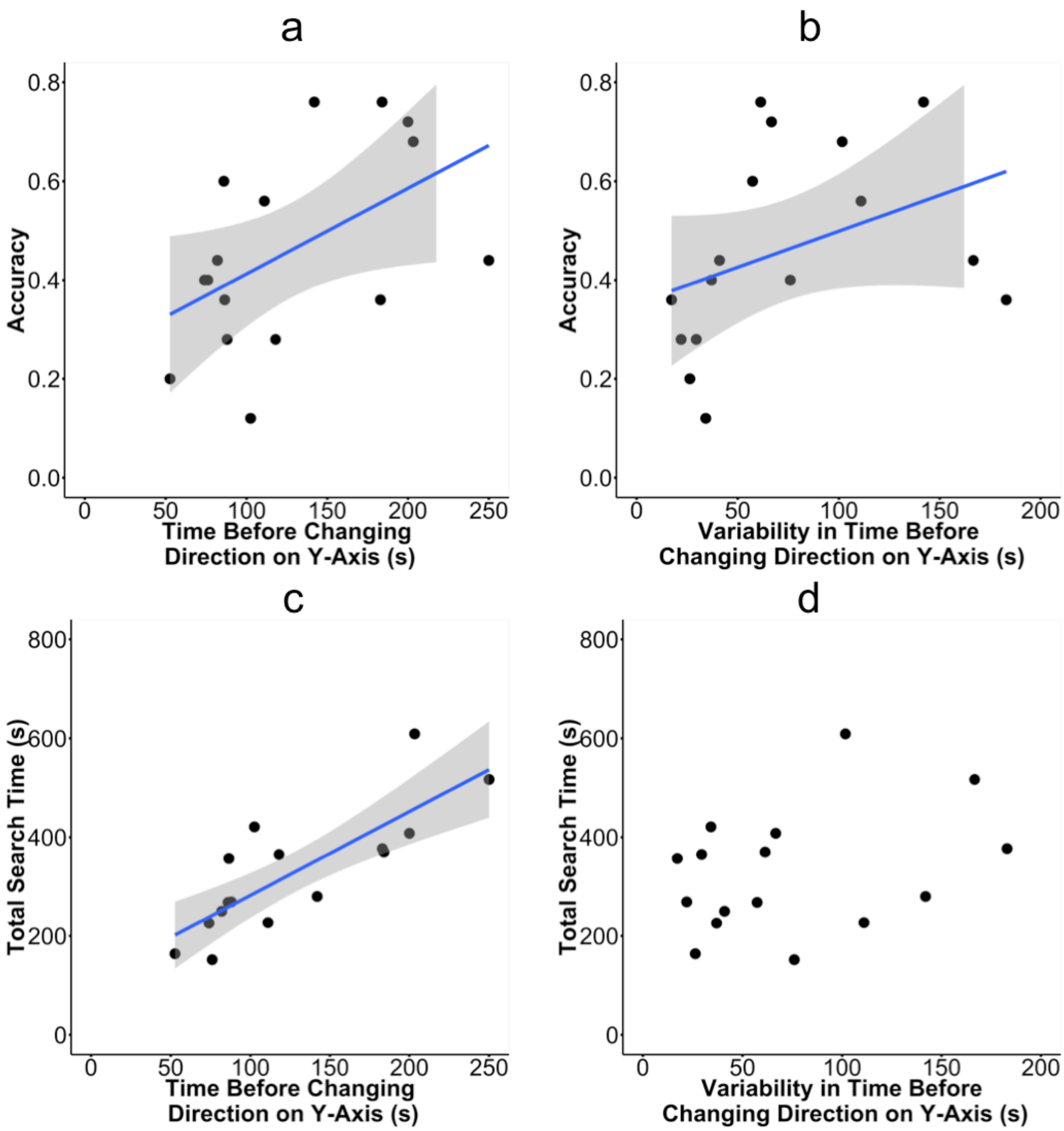

Figure 3.7 Search speed to predict Accuracy and Total Search Time for top-to-bottom-participants on the Y-axis. On the basis of the Fourier analysis, an approximation of participants' modal time before changing direction in seconds (s) on each axis was computed, using their dominant frequency component. By dividing the modal time before changing direction by the next most commonly occurring time before changing direction, the variability in time before changing direction in seconds (s) was computed. Accuracy was the proportion of coins detected and Total Search Time indicates the total amount of time spent searching in seconds (s). Figure 3.7 shows the modal time before changing direction (a) and the variability in time before changing direction (b) to predict Accuracy. Plus the modal time before changing direction (c) and the variability in time before changing direction (d) to predict Total Search time. Note, lines of fit are shown when a statistically significant relationship is present.

Discussion

In Experiment 1a, how searching for an unknown number of coins in an open space of grassland terrain is conducted was examined. It was predicted that accurate search would be reliant on (1) a regular, systematic search strategy following a structured path (Dalmaijer et al., 2015; Donnelly et al., 1999; Gilchrist & Harvey, 2006; Keech & Resca, 2010), (2) narrow search corridors (Koopman, 1943, 1980) and (3) slow, consistently paced movement (Koopman, 1943, 1980).

The basic behavioural data showed the task to be extremely challenging, with task accuracy being at 45%. Many targets were missed despite participants taking a significant amount of time to explore the search area (on average 7 minutes, 33 seconds). Targets were, however, detected at a fixed rate throughout search. Furthermore, accuracy was predicted by the confidence ratings given by the participants, suggesting participants had some idea of how accurate they had been in the task.

The route maps showed participants tended to search using an 'S' shaped strategy, though sometimes embedded within a more complex pattern (see

Figure 3.3). This was confirmed in the analysis of data extracted using the Fourier analysis. Given that all participants exhibited regularity in their search path, it was not possible to explore whether a regular search path predicted higher accuracy or not. However, the regularity of the 'S' shaped path made it possible for participants to define a search path that rarely contained crossovers and where the search width varied between 1 and 2m wide. Following an 'S' shaped path minimized the memory demands inherent in the task relative to if a more irregular path was followed (Gilchrist et al., 2001). Presumably this width of search corridor adopted by participants was set according to their beliefs about the salience of targets in the context of the environment in which they were being sought. More or less salient targets would, respectively, lead to use of a wider or narrower search corridor.

Given that participants opted to search using a common search width, an important question is, how exhaustive is each participants' search within their search corridor? Accuracy varied markedly across participants despite using the common search width and so using the common search width did not ensure that search was exhaustive. At least for some participants, the failure to search exhaustively was associated with a reduced modal time before turning. The implication of this speeded search is that areas of the search corridor were left unexplored as forward body motion occurred at a rate too fast for the sweep of left-to-right head movements and associated eye movements. The fact that accuracy was associated with confidence is consistent with participants having some insight into the likelihood of the success or failure of their attempt at an exhaustive search (metacognitive awareness).

One might think of this as a simple speed-accuracy trade-off but its observation is important. An observer tasked with ensuring or judging the quality of a search is unlikely to be able to make such a determination from the search path but it may follow from measuring differences in the time taken for search. Interestingly, slowed search distinguishes experts from novices in airport baggage screening tasks (Biggs et al., 2013). Calibrating how long a search task requires is a skill to be learnt both in complex visual searches and searches for targets placed in a more complex physical environment.

Variable search speed was also associated with increased accuracy, rather than a consistent pace, as predicted. This variability of search speed and accuracy found for the top-to-bottom participants was unexpected. On reflection, however, it is likely to be an effect associated with the task itself. Careful searchers slowed to ensure targets were clearly identified and marked as detected by the experimenters leading to variability in search speed as being identified with increased accuracy.

It is possible that the failure to find a relationships in the left-to-right group for time before changing direction and variability in time before changing direction for accuracy may be accounted for by the shorter time and distance between turns that participants had to make when moving left-to-right than top-to-bottom. Given that participants searching top-to-bottom were more accurate when they took longer before changing direction, the shorter distance before having to turn when searching left-to-right may have led to less efficient search. Measures based on participants movement through space may require sufficient movement time along axes, unfettered by the noise introduced by the slowing and speeding of turning itself, to become reliable indices of performance. In other words, the failure to find a relationship for accuracy for the participants moving left-to-right is likely to be a form of signal-to-noise problem.

Experiment 1b

One concern is about the generality of the conclusions that can be drawn from Experiment 1a. It is possible that target conditions used may have forced a specific search strategy where participants searched along an ‘S’ shaped path for reasonably densely packed targets. The data were generated in response to a single grid and a fixed set of 25 targets. Within the limits of the search area, these 25 targets had a specific configuration and density. Many studies have shown that target prevalence influences the conduct of visual search (e.g. Fleck & Mitroff, 2007; Godwin et al., 2010, 2015, Wolfe et al., 2007, 2005) and variations in target density are known to influence foraging (Cain et al., 2012; Wolfe, 2013), specifically, an increased number of turns in response to areas of higher density (Hills et al., 2013; Kalff et al., 2010). Might it be that the definition of search path and the width of the search corridors being searched are subject to change as target prevalence (along with target density and target configuration) varies? In Experiment 1b, Experiment 1a was repeated but with participants searching three different search grids, each with a different number of targets present.

Method

Participants

Fifteen participants (6 males and 9 females, mean age=24.67 years, SD=3.42) recruited from the University of Southampton community, with normal or corrected-to-normal colour vision took part in the study for course credit. As in Experiment 1a, participants were screened to ensure visual acuity and normal colour vision using Snellen (1862) chart for visual acuity and the Ishihara (1917) colour plates.

Apparatus

The apparatus was the same as in Experiment 1a.

Stimuli

The grid size was the same as in Experiment 1a. In Experiment 1b, three grids were used with three different numbers of targets (5, 15 and 25 targets). As target prevalence changed so did target density. Target density was, therefore, set at 1 target per 15m^2, 1 target per 5m^2 and 1 target per 3m^2. Experiment 1b was run in a different location and season to Experiment 1a and with a different depth of grass.

Procedure

The procedure was the same as in Experiment 1a, except all participants searched each of the three grid conditions. Condition order was controlled using a Latin Square design.

Results

The data from Experiment 1b were analysed in a similar manner to Experiment 1a in respect of accuracy, total search time and confidence. The focus of Experiment 1b was the influence of target prevalence on the search path used by participants. It was not designed to explore variations in movement speed as target prevalence is likely to influence the distribution of movement speed and slowing associated with the detection of targets as reported in Experiment 1a. Analyses of movement speeds are, therefore, not reported. In Experiment 1b, all participants detected coins, and so no participants were removed from analysis.

Behavioural Data

Basic measures of accuracy and total search time are presented in Table 3.1. On average, participants reported a confidence score of 7.13 on a ten-point scale for the 5 target condition, 7.87 for the 15 target condition and 8.47 for the 25 target condition. Accuracy, total search time and confidence were compared across the three target prevalence conditions. A series of one-way ANOVAs revealed no effect of Condition for accuracy (F(1, 14) = .392, p = .542, ges = .027) or total search time (F(1, 14) = .141, p = .713, ges = .01). Condition did reach significance for confidence (F(1, 14) = 32.941, p = <.001, ges = .702), with participants being more confident searching the 25 target condition than the 5 target condition. Regressions exploring the relationship between confidence and accuracy in each of the target frequency conditions showed none to reach significance (*ps* > .807).

As in Experiment 1a, for each participant in each condition the time of finding each target was regressed against the ordinal number of that target as found by the participant. (i.e. 1st, 2nd, 3rd etc.). The range of adjusted R-squared values were .733

to .995 (all *ps* <.238, the linear fits for 5 participants failed to reach significance), .84 to .992 (all *ps* <.002) and .923 to .993 (all *ps* <.001) for target prevalence of 5, 15 and 25 respectively. On average the first target was found after 35, 7 and 5in the 5, 15 and 25 target prevalence conditions respectively. On average, the last target was found 42, 17 and 13 seconds before terminating search in the 5, 15 and 25 target prevalence conditions respectively.

Search Strategy

Of the 15 participants, technical difficulties led to two corrupted files such that complete data was available for 13 participants. Example route maps are shown in Figure 3.8. Visual inspection of these route maps confirms the accuracy and search time data presented above, search strategy is mostly consistent across conditions. Of these 13 participants, 7 searched top-to-bottom in all conditions, 4 searched left-to-right in all conditions (2 searched with a hybrid strategy). For these participants there is no evidence of the route of the search paths changing with differences in target frequency.

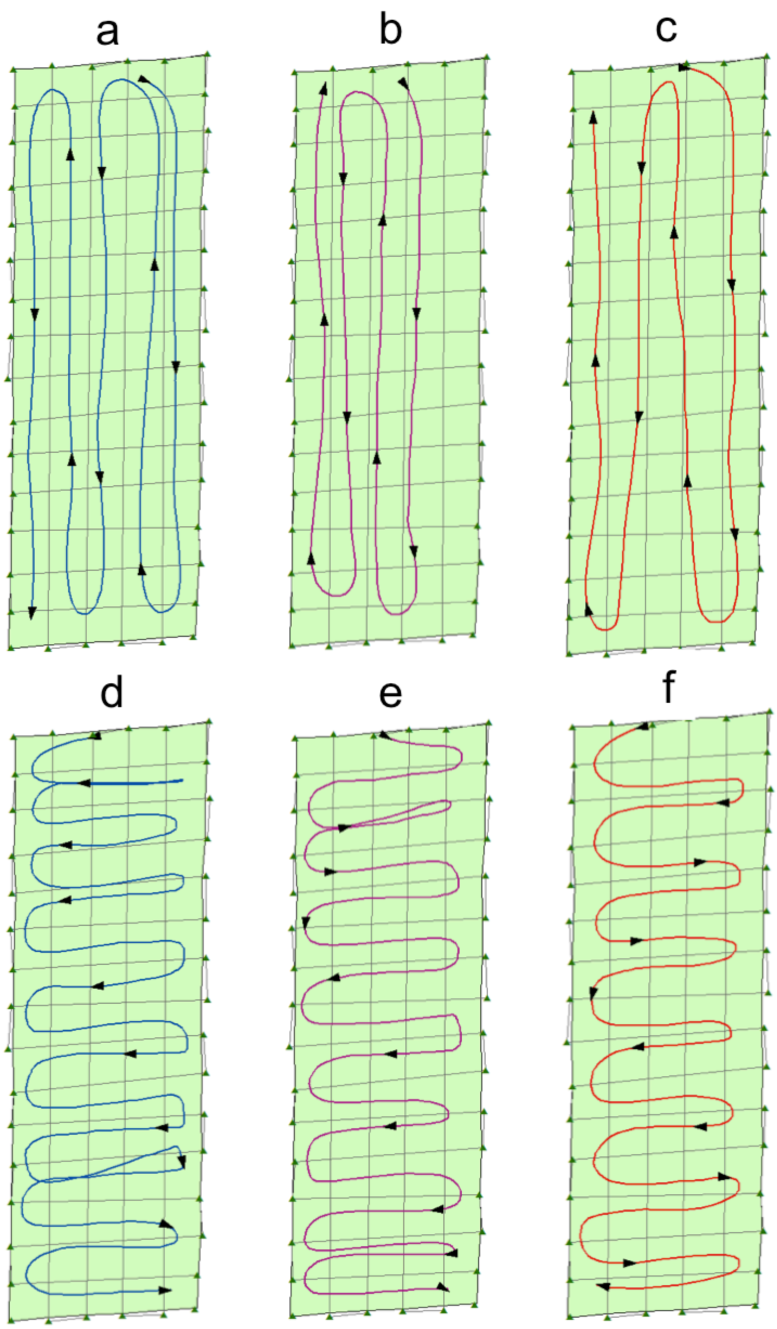

Figure 3.8 Route maps for Experiment 1b. Illustrations of examples of routes taken when searching. A participant using a common search strategy of top-to-bottom across target prevalence conditions (a, b, c) and a participant using a common search strategy of left-to-right across target prevalence conditions (d, e, f).

Discussion

Experiment 1b was easier for participants than Experiment 1a. The conditions of the experiment (both physical conditions on the ground, and the fact that participants were better practiced) led to higher accuracy in Experiment 1b than 1a.

Despite these differences, participants invariably adopted one strategy and then maintained it across conditions. The strategy chosen by participants was identical to that reported in Experiment 1a: an 'S' shaped path with a marginally higher likelihood for paths to go from top-to-bottom than left-to-right. A strategy that, irrespective of target prevalence, led targets to be detected at a consistent rate throughout the period of searching and with similar levels of success[2]. These data are consistent a view that, at least within the limits of the target prevalence tested in the present experiment, the search paths for searching for targets in open ground are determined in a manner independent of target prevalence.

While the search path is determined independent of target prevalence, the confidence with which search is conducted was not independent of target prevalence. Confidence in search reduced with reduced target prevalence. While participants were not told how many targets were in each grid, confidence may have been lowered by the experience of finding fewer targets in the 5 target condition relative to the experience of finding an increased number of targets in the 25 target condition. Whatever the underlying cause of the reduced confidence in search when target prevalence was low, an important point is that, as in Experiment 1a, the lowered confidence of participants did not change the search path. Furthermore, Experiment 1b confirms that the search path followed in Experiment 1a was not a result of the relatively densely packed targets but holds over much more sparse target densities.

Experiment 2

The coin detection task conducted in Experiments 1a and 1b was motivated by tasks such as police searching for clues to crimes in open grassland and soldiers searching for IEDs in theatre. In everyday life, such tasks are rarely performed by individuals working alone, but instead by searchers working in teams. Two strategies for team working can be used. Search can be split across the search space, or checked through sequential independent searches. Splitting search across participants might

[2] It is true that a number of participants in the 5 target prevalence condition did not have a significant linear fit between the number of the target detected and time. However, there are many fewer data points in this condition than in the 15 and 25 target prevalence conditions. As such, less reliability should be expected. Regardless, the general trend of these data is linear when considering all participants.

seem obviously advantageous. Doing so can minimize load allowing faster searching to occur (Brennan et al., 2008).

However, the evidence from Experiments 1a and 1b suggests that the approach to searching across participants might not translate into improved search accuracy when splitting search across space. The failure to find a relationship between width of the search corridor and accuracy implies that being seen to adopt an 'S' shaped search path should not translate into beliefs about equivalent search competence. Put another way, splitting search by area allows the risk of a poor searcher appearing competent but working unchecked, and a risk of coordinating search effort (Brennan et al., 2008).

A better strategy might be for pairs of searchers (henceforth dyads) to benefit from the independent probability summation that occurs when searching independently across search spaces. Independent probability summation is the summed probability of detecting each target that comes from the addition of each searcher's likelihood.

In Experiment 2, Experiment 1a was repeated but with dyads searching for coin targets. They were allowed to define their own search strategy and were free to communicate in whatever way they felt appropriate. Dyads electing to split search by area and those conducting independent search were defined post-hoc. The performance of pairs electing to split search by area and those electing to not do so was compared, as was the performance of both groups to the individuals reported in Experiment 1. It was predicted that if individuals in the dyad worked independently of each other (i.e., search effort doubled over the entire space), then accuracy would be higher as performance would benefit from the summation of two searchers, without the risk of coordination attempts. This increased accuracy will come at the cost of increased time relative to those who split search by area. Finally, it was also predicted that as in Experiment 1a and 1b, participants would search systematically using an 'S' shaped route searching primarily either top-to-bottom or left-to-right.

Method

Participants

Thirty-four participants (14 males and 20 females, mean age=26.38 years, SD=5.79) recruited from the University of Southampton community, with normal or corrected-to-normal colour vision took part in the study for course credit. As in Experiment 1, participants were screened to ensure visual acuity and normal colour vision using Snellen (1862) chart for visual acuity and the Ishihara (1917) colour plates. Participants were allocated into 17 dyads.

Apparatus

To ensure comparability across Experiment 1a and 2, both members of the dyad held a staff as they searched. The Total Station theodolite was, however, limited to recording coordinates for one participant in each dyad. Attempts were made to record similar data from the second member of each dyad using a satellite-based system but this proved unreliable and so these data were unfortunately not viable for the analyses.

Stimuli

The experiment was conducted on the same areas of grassland as in Experiment 1a and at the same time of year. The grid size was also the same, and the same twenty-five UK sterling two pence coins were used in the same grid cell positions.

Procedure

The procedure was the same as in Experiment 1a, but with participants working in dyads. The Dyads were instructed to work together in the task, but no instruction or guidance was given to assist with determining a strategy for the dyads to work together. Dyads were required to give a combined confidence rating, i.e. they were asked to agree on a score on the ten-point scale.

Results

In Experiment 2, all dyads detected coins, and so no dyads were removed from analysis.

Behavioural Data

Considering dyads as a single group, basic measures of accuracy (total for the dyad) and total search time are presented in Table 3.2, alongside the individual's data from Experiments 1a and 1b. On average, dyads detected 68% of targets and spent just over 6 minutes searching. On average, participants reported a confidence score of 7.12 on a ten-point scale. To assess the advantage of dyad search over individual search, the data from Experiment 2 were compared to those from Experiment 1. Independent t tests revealed dyads searched more accurately than individuals $t(42.757) = 4.089$, $p < .001$) but that there was no difference in search time ($t(34.374) = -.89$, $p = .19$). Note, for the dyads, search time is defined as the time at which both members of the dyad agreed search had finished, rather than the summed time of the two searchers. Dyads were more confident in their search performance than individuals $t(36.438) = 1.75$, $p = .04$) though unlike the individuals in Experiment 1a, confidence ratings for the dyads failed to predict accuracy ($\beta_1 = .043$, $F(1,27) = 2.291$, $p = .151$, adj $R^2 = .07$).

Shared Searching

To explore how participants in each dyad split their search, for each dyad, targets were classified as detected by (1) participant one, (2) participant two, (3) both participants, or, (4) missed by both participants (see Table 3.2 for means).

Table 3.2 Shared searching data. The Mean number of targets detected by only one participant in a dyad (Sole Hits), the number of targets detected by both participants in a dyad (Joint Hits) and the number of targets missed by both participants in a dyad (Misses). Note, parentheses indicate SD.

Sole Hits	Joint Hits	Misses
12.29 (4.81)	5.35 (5.56)	7.35 (2.8)

These data were used to classify, for each dyad, the number of misses, individual hits, and joint hits. The number of joint hits was subtracted from the sum of the number of individual hits from each dyad, and divided by the total number of hits. This gives a ratio from +1 (totally independent) to -1 (totally shared) to give a measure of target detection strategy that is independent of accuracy.

To test whether the search style adopted by the dyad influenced accuracy, the dyads were then split into two groups, (1) dyads with a split search ratio higher than 0, meaning they tended to search independently (n = 5) and (2) dyads with a split search ratio lower than 0, meaning they tended to split search between them (n = 12).

The data from the Theodolite system provided data to analyse the search strategy from one member of each dyad, and route maps were used to seek direct evidence that some dyads tended to split search. In theory, this gave us a total of 17 participants, however, due to technical faults two participants could not be included in the search strategy data, giving a total of 15 participants. Inspection of the route maps from the individual within each dyad for which movement data was recorded provides evidence consistent with splitting the dyads into shared and independent searchers (see Figure 3.9). These route maps show independent searchers tend to inspect the whole search area whereas split searchers tend to inspect a sub area of the search space. Of the 5 independent searchers, 2 searched top-to-bottom, 3 searched left to right and 1 performed a hybrid search strategy. Of the 9 split searchers, 4 searched top to bottom and 5 searched left-to-right.

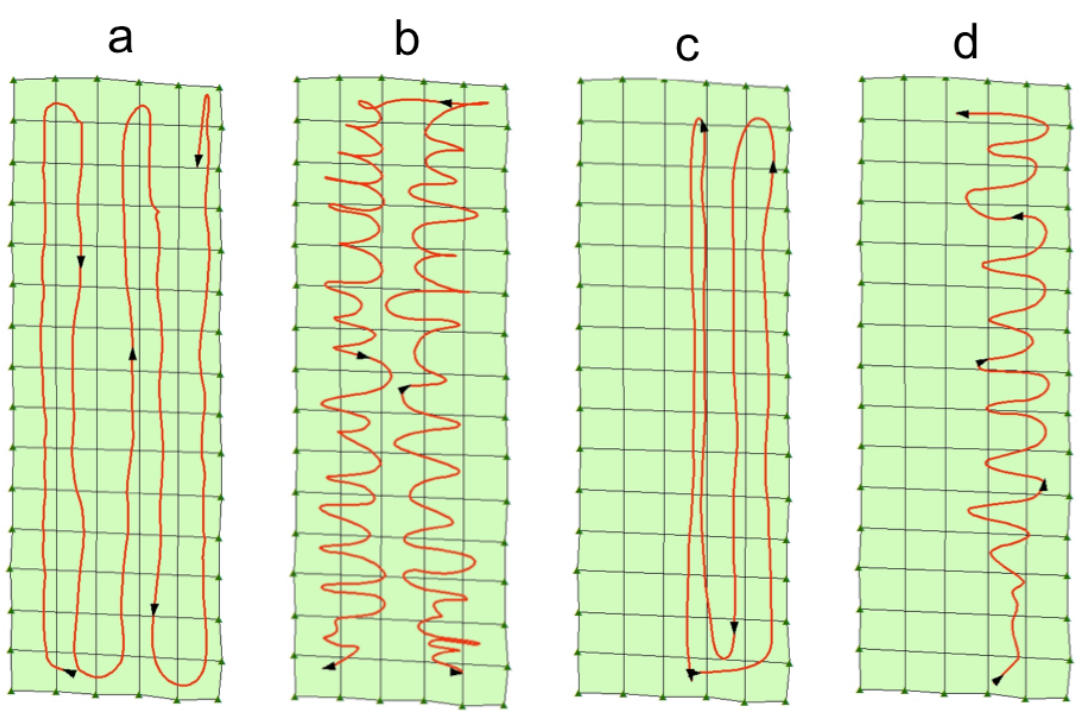

Figure 3.9 Route maps for Experiment 2. Illustrations of examples of routes taken by independent searchers searching top-to-bottom (a) or left-to-right (c) and split searchers searching top-to-bottom (c) or left-to-right (c).

Both dyad groups, whether tending to search independently or splitting search, were significantly more accurate than individual searchers (t(11.479) = 5.36, p < .001 and t(35.087) = 2.885, p = .006 respectively). A final t-test confirmed that dyads tending to search independently were more accurate than those tending to split search (t(8.699) = 3.053, p = .014). Dyads who tended to split search took significantly less time to search than individuals (t(37.824) = -2.83, p = .007) but there was no significant difference between dyads who tended to search independently and individuals (t(5.162) = 1.079, p = .328). The difference in search time between dyads who tended to search independently and dyads who tended to split search failed to reach significance (t(4.389) = 2.315, p = .076).

Search Strategy

The movement data recorded from the individual within each dyad does allow us to explore one other issue (note that computation of search width and search speed makes no sense as the data come from only one member of each dyad). As in Experiments 1a and 1b, each participant can be classified as searching primarily top-to-bottom or left-to-right, or using a hybrid strategy (see Figure 3.9). 6 participants used a top-to-bottom strategy (2 independent and 4 split) and 8 (3 independent and 5 split) used and left-to-right strategy. Using a Fisher's exact test, and comparing these data against those from Experiment 1 (where across 1a and 1b, 23 used a top-to-bottom strategy and 16 used a left-to-right strategy), revealed a non significant result (p =.358). Working in dyads did not significantly change the likelihood of direction of searching.

Discussion

The results of Experiment 2 confirm the results of Experiment 1a in showing the coin detection task to be very demanding. While accuracy was higher in Experiment 2 than 1a, performance was still not at ceiling for any pair. More importantly, Experiment 2 showed two critical results. First, dyads searched more accurately than individuals searching the whole space alone. They did this whichever strategy they adopted, whether splitting search or searching independently of each other. Whatever the cost of coordinating search effort (Brennan et al., 2008) across the open area of grassland, these costs are overcome by dyads. The benefit to search accuracy of dyadic relative to individual searching comes with no cost or benefit to search time (when both dyad groups are considered together). Within the same unit of time available, dyadic search benefitted from twice the resource being applied to search. Second, and critically, as predicted search performance was best when dyads searched independently rather than when splitting search. Dyad accuracy benefited from the summing of the performance of two independent searchers rather than splitting search. Furthermore, as in Experiment 1a and 1b, and from the data available, it seems participants searched systematically using an 'S' shaped route searching primarily either top-to-bottom or left-to-right.

It is concluded that, at least for the experimental conditions used in Experiments 1a and 2, search is more accurate when carried out by dyads over individuals, but that dyads are most accurate when providing two independent passes over the open space rather than when trying to split search such that overall performance benefits from independent probability summation.

General Discussion

Chapter 3 set out to explore how participants search for targets on open grassland. The task reflects a class of search problem that has received little attention in psychology, but is one which is a common problem in a wide range of security-critical and evidence-gathering scenarios. Participants searched open grassland, either individually (Experiments 1a and b) or in dyads (Experiment 2), for an unknown number of small targets of low salience (coins). Significant numbers of misses were made across all experiments indicating that although possible, all versions of the task were difficult.

Experiment 1a revealed individuals typically use an 'S' shaped search strategy with a common search width of between 1 and 2m, and with targets being found throughout the search time. Furthermore, when searching from top-to-bottom, increased accuracy is associated with slow search (i.e., longer time before changing direction) but with variable speeds (i.e., stopping to check targets). For all participants, targets were detected at an even rate throughout searching confirming that task performance was not affected by reducing vigilance over the time course of searching or by search being satisfied after finding targets early in search. The "S" shaped strategy minimizes memory load and maximizes the possibility of an exhaustive search (Gilchrist et al., 2001). It may be the case that the 'S' shaped strategy is a function of the shape and area of the search region. While the $75m^2$ search area was much larger than that typically used in visual search studies, it was still a regular rectangle and defined in a manner that allowed planning (i.e., could be seen as one, and defined by cones). Further studies could explore how search strategy is influenced by the global shape and size of the search area.

Experiment 1b revealed that the accuracy, total search time, 'S' shaped strategy and the even rate of target detection across the duration of searching are unchanged by varying target prevalence. This evidence is viewed as important in as much as it shows that the specific conditions of Experiment 1a did not lead to an artefactual result, but instead one that generalizes beyond one set of conditions. It is, of course, the case that in Experiment 1b the same participants searched in all conditions. As such, they will have experienced search for targets occurring at different prevalence rates in a reasonably short period of time. For this reason, Chapter 3 does not claim that these data definitively demonstrate that target prevalence does not influence search strategy in this task. For that conclusion to hold, more extensive experimentation would need to be conducted, perhaps using a blocked design and even lower levels of target prevalence.

The existence of a structured approach to searching for coins is important. Structured search allows the possibility individuals within dyads might coordinate search strategies. Coordinating search by splitting search across space might have reduced total search time if not also improved accuracy. In fact, Experiment 2 revealed dyads were more accurate than those searching individually, even when splitting search over space.

Experiment 2 showed that the most effective strategy for dyads was to conduct two independent searches rather than splitting search. It is important to note that conducting independent searches is not a failure to coordinate search across dyads. It is possible that dyads held an implicit understanding that there are risks and costs associated with the coordination of split search (i.e. even if a partner's search appears to be conducted in a serious manner, it does not necessarily predict good target detection). These costs lead to worse performance than can be achieved through summing the independent probabilities for target detection. The clear conclusion when searching for threat critical targets is that searching for targets in open ground is enhanced by dyads working in tandem but searching independently. Of course, how this conclusions stands up as search teams that go beyond dyads, and as the search area increases beyond that used in the present experiments, is uncertain. Given that the performance of the dyads was best when they searched independently, it is also important to consider the difference between two members of a dyad searching independently of each other, compared to two searchers searching the grid independently at different times. Even searching independently, it is likely that the dyads in Experiment 2 coordinated somewhat to agree on such a strategy. The presence of a second searcher likely held a social influence over participants, and influenced factors such as where they began their search, their speed of search, and when they finished their search. Furthermore, it may be that participants took less care when searching the half of the grid already searched by the other member of the dyad. Further experimentation could explore this.

It is important to note that in the present study, the tendency to opt for either a split search or independent search strategy reflected a decision made within the dyad. Decisions about strategy might be open to review throughout the course of conducting search tasks if, for example, a partner appeared to be searching without due care and attention. There is evidence that humans can make attributions about others from observing motor control during task performance (Wolpert, Doya, & Kawato, 2003). The data from Experiment 1a do suggest, however, that searchers should be wary of believing search is being conducted accurately simply by observing the strategy adopted by partners and the time taken to search.

In a related manner, the relative influence of social facilitation (Triplett, 1898; see also Strauss, 2002) and social loafing (Karau & Williams, 1993) may also raise or diminish overall performance. In the case of the present study, comparison across Experiments 1 and 2 shows social facilitation tended to improve performance by raising accuracy without changing search time. Of course, what influences decisions within the dyad to work together in one mode or another may reflect multiple factors. Low levels of willingness to trust and high levels of thoroughness may both be reflected in high levels of independent searching of the whole search space. There is evidence that differences in levels of individuals extraversion and agreeableness, and overall conscientiousness of pairs influences the likelihood of individuals attributing challenges to paired working (Bono, Boles, Judge, & Lauver, 2002). The present study does not provide data to help understand the role of these mediating variables in predicting dyadic search strategy. Nevertheless, it should be acknowledged that extraversion, agreeableness and conscientiousness may predict the likelihood of individuals conducting a systematic search across open space, and the strategy of search pairs. It is an issue worthy of future study.

The nature of the search task is also likely to influence the decision making and performance of the dyads. Although the present task models the detection of clues in crime scenes and IED detections in war zones, it was not safety critical and therefore did not have the same heightened emotions that are likely to occur with a safety critical task. The gravity of the consequence of a miss or a failure to search exhaustively and efficiently may modulate the strategy dyads adopt in coordinating. For example, if the dyad opt to conduct two independent searches then they may increase their accuracy, but they double their individual risk in terms of the extra ground they cover with increased exposure to threat.

In conclusion, Experiments 1a revealed the importance of systematic search for target detection, with participants using an 'S' shaped search route with a common width of search corridor. Participants found targets at a common rate throughout searching along the search route. When searching top-to-bottom, increased accuracy was associated with slower searching. Accurate target detection occurred when time allowed for eye, head and body movements to be made to search across the width of the search corridor. Experiment 1b revealed that variations in target prevalence did not change the shape of the search route followed by participants, nor the overall accuracy or time spent searching. It did, however influence the confidence in their search performance with confidence lowered when targets were relatively infrequent. Experiment 2 showed how dyads improve target detection relative to when search is conducted by individuals. This is true when splitting search but especially when improved accuracy results from the summed total of conducting two independent searches.

Chapter 3 has outlined a number of important findings, methods and analyses in relation to searching for targets in open space. Future studies should seek to show how the search strategies that have been outlined generalize or are modified by stimulus and environmental conditions. For example, whether systematicity survives under conditions of very low target prevalence (Godwin et al., 2010) or target absence (Schwark et al., 2013), changes in the size and shape of the search area (Smith, Gilchrist, et al., 2008), variations in target type and identity (Menneer et al., 2007) and the presence of distraction and the possibility of concealment (Godwin et al., 2015). Furthermore, how dyadic working influences the effectiveness of these strategies in enhancing target detection for multiple targets placed in open ground. In particular, how the dyadic performance is changed by explicit instruction to follow specific strategies, time pressure and perceived risk.

Chapter 4

Rummage Search of a House: Comparing Novice Individuals, Novice Dyads and Expert Searchers

Some of the data from this chapter has been previously reported in two technical reports for the Defence Science and Technology Laboratory (Donnelly, Riggs, Mann, Godwin, Menneer & Liversedge, 2013).

In terms of the complex search task of rummage searching for objects in a house, Chapter 1 revealed a number of factors to be important. In addition to the issues explored in Chapters 2 and 3, Chapter 4 builds on the previous tasks in the present thesis with the addition of physical motoric search of objects which may be hidden or partially hidden, may overlap or may be broken down into parts. As with previous chapters, the implication of paired searching is also of importance, as expert searchers often carry out rummage search in pairs within larger search teams.

Rummage search (RS) is a form of real-world visual search where searchers engage in the visual and haptic search for targets. RS often requires searching over large areas such as a queue of vehicles, a number of houses or big public venues. Targets in RS may be visible and in plain sight, or they may be hidden from sight underneath or inside other object(s) such that some motoric action is required to find them. We engage in forms of RS in everyday life, for example when looking for our keys in a coat pocket or a mobile phone on a messy desk. However, when in security critical scenarios, RS is often conducted to find items that are purposely hidden. The police, military and border agencies perform RS to find weapons, drugs and IEDs. Despite being a skill that is trained, there have been relatively few studies exploring aspects of RS (Foulsham et al., 2014; Gilchrist et al., 2001; Jiang et al., 2014; Smith, Hood, et al., 2008). These studies do not provide much of an evidence-base for either understanding RS or improving its training. The primary goals of Chapter 4 are to present a framework for studying RS and to then use this framework to establish a clearer evidence base for determining how the training of novice rummage searchers might be enhanced.

Laboratory experiments on visual search are a starting point in providing some insight into RS. There are, however, a number of key differences between experiments investigating visual and rummage search, which make it difficult to develop an understanding of RS exclusively from visual search. Most importantly, visual search experiments are presented on a computer screen (see Chan & Hayward, 2013; Eckstein, 2011 for a recent review), and typically use simplified static stimuli such as coloured shapes or lines (Treisman & Gelade, 1980), where targets and distractors do not overlap (though see Solman, Cheyne, & Smilek, 2012). In contrast, RS search is large-scale, requires both visual and haptic search, and targets can be both complex, and hidden or partially obscured.

The handful of studies that have explored aspects of RS have shown some striking results that stand in contrast to visual search behaviour observed in the laboratory. First, while objects are often revisited in visual search tasks (Posner & Cohen, 1984; see also Klein, 2000), this is rarely the case in search where participants have to walk through the environment to uncover targets and distractors (Gilchrist et al., 2001; Smith, Hood, et al., 2008). Second, while target salience is a cue known to be important for increasing the speed of visual search (Itti & Koch, 2000, 2001), this has been found to only be true in RS if participants are primed to highly salient cues (Foulsham et al., 2014). Third, when free to walk around a search space, searchers are able to learn regularities in the environment (i.e. where a target is likely to appear) even across changes in viewpoint or perspective (Jiang et al., 2014); a finding which differs from visual search (Jiang, Swallow, & Capistrano, 2013). These results demonstrate that searchers who have to make movements beyond those of just their eyes to find targets make few revisits, suggesting searchers must employ a structure to their chosen search path. They also demonstrate that searchers require target salience to be primed in order to be effective, suggesting searchers must rely on something other than visual guidance (such as a structured path) to aid their search. Finally, searchers can learn regularities in the environment, which would aid the forming of a structured path. These conclusions are important as they show a first attribute of effective RS is planning and following a search path in a systematic manner.

Invariably the area around the search path taken during RS will not be defined solely by where the eyes can move. Head and body movements must also be made to ensure all possible target areas are searched. Using a term taken from studies of search and rescue, the width of the area around an individual that can be searched effectively given a set of eye, head and body movements is called the Effective Search Width (ESW; Koopman, 1946, 1980). In the general case, the technical limits of sensors (human vision, radar etc.), determines the width of the ESW. Targets that sit outside of the limits of the ESW will not, by definition, be detected. In RS, the ESW will be set by the need to coordinate eye, head and body movements to exhaustively explore the search area. The second attribute important for effective RS, therefore, is that the ESW is set to ensure all potential target locations can be searched exhaustively.

The need for RS to go beyond a purely visually guided target search is a consequence of the fact that at any location targets might not be in plain view (i.e. targets might be occluded or hidden). In these circumstances, rummage searchers must identify not just targets but the affordance of the environment to hide targets, and then determine and implement actions to counteract these affordances. For example, targets might be hidden behind objects that need to be moved, or placed inside other objects that need to be opened or compressed (e.g., cupboards or cushions). The requirement to encode the environment for actions that might reveal targets extends the concept of target templates in visual search. The templates guiding RS must define targets but also some abstract conceptualisation of the visual attributes of spaces that might hide targets. As such, the third attribute of effective RS is to be able to encode the affordances of environments to hide objects as well as to enact the actions required to counteract these affordances. These three attributes are considered here to be core characteristics of RS, and suggest that for RS to be effective, searchers need to consider each attribute search in a manner that encompasses each of them.

Chapter 4 is interested in the type of RS that occurs when searching for evidence related to a crime. In such situations, RS must be exhaustive as the risks associated with failing to find *all* possible targets are high. When the task is to find clues to a murder, failing to find all targets can lead to failed convictions; when the task is to find IEDs, failure to find all targets can lead to injury and even death. Exhaustive RS requires faultlessly performing in accordance with each of the three attributes listed above.

Whether rummage searchers can ensure all targets and potential target locations are attended and explored is the question that is fundamental to the present study. There are no data currently available that speak to this issue. Returning to consideration of visual search, what we do know is that the capacity to conduct exhaustive visual search for targets is rather poor. In the case of visual search for rare or multiple targets, a number of studies have shown preventing an early termination of visual search to be a real challenge. For example, as discussed in Chapter 1, searching is known to be incomplete for infrequently occurring targets (Godwin et al., 2010; 2015; Wolfe et al., 2007; 2005), searching for more than one instance of a target (Cain & Mitroff, 2013; Tuddenham, 1962), searching for two target types simultaneously (Menneer et al., 2007), or searching for an unknown number of targets (Cain et al., 2012; Wolfe, 2013). While these data from visual search studies might not be directly relevant to RS, they certainly raise the concern that participants conducting RS may not be exhaustive in finding potential target locations, or searching those potential target locations effectively.

In the present study, participants engaged in RS of a residential house for incriminating evidence about a crime. Participants were given a briefing and asked to search for targets relevant to the brief. They were instructed to locate any firearms or money associated with a robbery at a bank, and any drugs that may be present at the property too. Participants did not know the number, location or distribution of potential targets. There were, in fact, six targets that varied in type and size (a small and large gun, a small and large roll of money and a small and large quantity of 'drugs'). One target was plainly visible, some could be found after moving or opening objects and some required careful haptic examination in order to be found. Borrowing a term from eye movement research, areas that could hold targets were defined as regions of interest (ROIs).

The exhaustiveness of RS was measured with respect to the coding of the house into separate ROIs and operationalized in terms of three possible outcomes. Participants might engage in; first, a visual inspection of visible areas (i.e., a table top); second, visual inspection that would be improved by moving an object (e.g. an ornament) that might act as an occluder or opening (e.g., a cupboard door) objects; and third, haptic examination of spaces (e.g., the inside of a hidden cupboard) or objects (e.g., the compressing of a cushion) where visual inspection alone would have been incomplete. RS behaviour was examined by response accuracy, search time, a series of eye movement measures and measures of haptic action. The use of mobile eye movement technology and the accompanying measures allowed us to determine which ROIs had been attended, and the actions made in respect of those ROIs.

The methodological approach, response taxonomy and performance measures taken in the present study emerged out of extensive pilot testing performed in conjunction with military trainers. In the current study, first pairs of expert rummage searchers (expert dyads) were tested. Expert rummage searchers typically work in pairs and therefore their normal working practice was not changed. Working in pairs requires them to double-check each other, for an exhaustive search of all possible targets locations. This is because of the requirement to ensure all possible targets are found when the consequence of failing to find targets is high. It is expected that both expert searchers would engage in an exhaustive search of ROIs, exhaustive search defined as attending to all possible target locations AND examining each of the attended locations to the extent required to determine the presence or absence of a target. These two indices of exhaustiveness are independent but experts were expected to score highly on each dimension of exhaustiveness.

To explore RS in an untrained population, a group of novice participants were also tested. These novice participants were either tested in pairs (novice dyads) like the experts, or as novice individuals searching alone. Although the experts are trained to work as a pair, it was not known whether this would help or hinder the novices. It seemed possible to us that novices may feel subject to less social pressure to terminate search if working alone than as part of a dyad. If so, novices may show their fundamental ability to search exhaustively better when searching alone than as part of a dyad. For a fair comparison with experts, both novice dyads and novice individuals were tested. It might be that both novice groups differ from the experts, but for the final conclusion in respect of the difference between expert and novice rummage searchers, it will be based on whichever novice group is the least different from the experts. It was expected that evidence would show experts to conduct exhaustive RS but novices would not. Whether the failure to search exhaustively for novices would be found for both measures of exhaustiveness was an open question. The simple hypothesis that was explored in Chapter 4 was that the failure of novices to RS exhaustively would occur through a failure to attend to all possible target locations and a failure to search attended locations completely. In other words, it was predicted that, in contrast to experts, novices would fixate fewer ROIs and explore these ROIs less exhaustively. Furthermore, it was predicted that participants would be less likely to be exhaustive as the motoric effort required to effectively search a ROI increased. For instance, participants would be more likely to carry out an exhaustive search on a ROI that required an object to be moved, than on a ROI that required full haptic examination. The results will be considered as evidence for validation of the RS framework if it helps to capture the differences between expert and novice searchers in a meaningful way. If the framework is validated then it will provide a taxonomy that would be useful to future researchers and an analysis helpful in developing training.

Method

Participants

Seventy-four participants (39 males and 35 females, mean age = 24.97 years, SD = 4.85) with normal or corrected-to-normal colour vision took part in the study. Sixteen participants formed eight expert military dyads and were volunteers from the Royal Logistic Corps who had completed search training. Participants were remunerated at the standard rate for military trials (approximately £15 in total). Forty-eight participants recruited from the University of Southampton community formed twenty-four novice dyads and took part in the study for either course credit or payment (approximately £12). Ten more participants from the University of Southampton searched individually as novice individuals. Participants were screened to ensure visual acuity and normal colour vision using Snellen (1862) chart for visual acuity and the Ishihara (1917) colour plates. The study was performed in accordance with the Declaration of Helsinki and was approved by the University of Southampton, School of Psychology ethics committee; informed consent was obtained from all participants.

Apparatus

Eye movements were recorded using an Ergoneers Dikablis wireless head mounted eye tracking system consisting of two small cameras (one directed toward the scene, one toward the eye) mounted upon a lightweight pair of spectacles. Both cameras were accurate to 0.5 degrees of visual angle, operating at 25 Hertz (Hz) per second, sampling the environment and eye every 40 milliseconds. The video footage was wirelessly transmitted to a pair of host laptops.

Stimuli

The experiment was conducted within four rooms (Kitchen, Living Room, Large Bedroom and Small Bedroom) of a fully furnished residential house (see Figure 4.1). The rooms differed in their structure, function, form and size. The rooms were treated as a single trial. There was no particular interest in differences between rooms and so they are not considered further.

Figure 4.1 The four rooms of the fully furnished residential house. The Kitchen (a); Living Room (b); Large Bedroom (c) and Small Bedroom (d).

ROIs were defined as items of furniture (such as the bed, the desk, the wardrobe), other individual items (e.g. paintings), or sets of items (e.g. shelves of books and records) etc. Wall and floor surfaces were not coded as ROIs.

ROIs were coded in terms of the actions required to search them exhaustively (visual inspection, visual inspection involving the movement of objects, or haptic examination following visual inspection (and possibly object movement)). Although a participant could engage in visual inspection alone, no ROI could be searched exhaustively by just visual inspection alone (i.e. all ROIs required at least moving or opening an object). Although one of the targets was detectable by visual inspection alone, that target was positioned on a defined ROI that also required the moving of objects for search to be exhaustive; e.g. a windowsill which had multiple items along it. A ROI *could* have required visual inspection alone, but the coding of the ROIs in the present study as items of furniture or sets of items meant that all ROI required at least some action to be searched exhaustively. For a full list of ROIs and the set of actions required to conduct a full RS, see Appendix A. The order in which the rooms were searched was controlled using a Latin Square design.

The targets (see Figure 4.2) were representative of a crime scene evidence detection scenario. The number of targets was varied across rooms, with three located in the Small Bedroom, two in the Living Room one in the Large Bedroom and zero in the Kitchen (see Table 4.1) Targets were placed in the same locations for all participants.

Figure 4.2 The targets used in the RS. A large and small weapon, a large and small stash of money and a large and small stash of 'drugs'.

Procedure

Following successful screening for visual function, participants were required to read through a short PowerPoint presentation explaining the RS task and crime scenario briefing[3]. During this presentation, photographs of four guns, four rolls of cash, and four packets of drugs were shown. These photographs were not the actual targets, but provided generic exemplars of the types of targets hidden in the house. Once fitted with the eye tracker, participants completed a nine-point calibration procedure.

Participants were informed that an armed robbery had recently taken place, and that they were to search for incriminating evidence. They were instructed to search each room until they (or the dyad) were confident they had completed their search of that room. Once they had completed the search of a given room, they were shown to the next room to be searched. They were not allowed to return to rooms once they had determined that a room had been searched thoroughly.

Participants working as individual searchers completed the task alone, and novice dyads were told of the importance of working together to search as accurately and efficiently as possible. No instruction or guidance was given to assist with determining a strategy for pairs to work together. Experts searched as per their standard instructions.

Results

The point of fixation was overlaid on the video of the room view from each participant and then the marked video was used for coding the data. The coding of each video was very time consuming and was carried out twice for each participant. The first pass through the video identified the ROIs fixated. The second pass through scored the action performed at each ROI fixated. It was not practicable or possible to carry out a double-blind coding procedure, as the identity of the participants (experts versus novices, dyads versus individuals) was clearly visible to the coder on the video

[3] While the basic presentation on RS was common across all participants, the naïve participants were given different threat briefs (that the target spotted by an eye witness was congruent in size with the threat brief (n = 8), incongruent in size with the threat brief, i.e. the threat brief was inaccurate (n = 8) or a more general threat brief not specifying a specific target at all (n = 8)) or some additional details about rummage searching regarding possible sources of error (a form of error management training) (n=12) or no additional details (n = 12). Statistical comparisons of performance speed, accuracy and exhaustive search, as well as underlying eye movement measures showed that those receiving the basic information and those with minor variants differed on only 1 out of 16 tests. For this reason, participants have been collapsed across these minor variations to increase the statistical power of the study.

stream. In order to validate the coding, two specially trained coders carried out all coding. The coding of dyads was split so that each coder coded an individual within each dyad. A sample of the annotated video footage has been made available; this can be viewed at https://dx.doi.org/10.6084/m9.figshare.3817380.v1.

The coding of the action(s) performed at each ROI fixated was used to examine how exhaustively participants searched ROIs. Exhaustiveness might require (1) visual inspection, (2) visual inspection following moving or opening an object(s), or (3) haptic examination (following moving or opening an object(s)). In practice, and due to how the environment was constructed (ROIs were defined as items of furniture and sets of items), there were no ROIs that could be searched exhaustively by just initial visual inspection alone, that is to say, all ROIs required at least an object to be moved, or a container to be opened for exhaustive search. Failure to perform the full set of actions required for exhaustive search was coded as an incomplete search. The actions required to search each ROI exhaustively are shown in Appendix A.

Behavioural Data

Response Accuracy

Mean target detection rate was 92% (SD=18%) for the expert dyads, 59% (SD=19%) for the novice dyads and 68% (SD=34%) for the individual searchers (see Table 4.1 for the accuracy rate broken down by target type). One sample *t*-tests, comparing accuracy rates against ceiling (100%), showed both the individual and novice dyads performed significantly below ceiling ($t(9) = 2.967$, $p = .016$; $t(23) = -10.552$, $p < .001$), but the expert dyads were at ceiling ($t(7) = -1.323$, $p = .228$). A one way ANOVA revealed a main effect of Group ($F(2, 39) = 6.003$, $p = .005$, ges = .235). Expert dyads were significantly more accurate than the novice dyads ($p = .004$). No other contrasts reached significance (*ps* > .119).

Table 4.1 Target Accuracy. The proportion of each participant group to detect each of the targets.

	Target Accuracy		
Target	Novice Individuals	Novice Dyads	Expert Dyads
Large Drugs	0.5	0.38	0.88
Small Drugs	0.5	0.17	0.75
Large Gun	0.8	0.88	1
Small Gun	0.8	0.92	1
Large Money	0.7	0.5	0.88
Small Money	0.9	0.71	1

Total Search Time

Search time was measured from the start of the first gaze on a ROI in a given room, to the end of the final gaze on a ROI in a given room. Participants indicated to the experimenters once they had completed their search of a room. With respect to dyads the end of search was measured when both searchers agreed that search was complete. Times for each room were summed to produce a total search time for RS of all rooms used in the house. The time for dyads is the search time for the pair (and not the time for each participant summed across both participants). Total search time was 01:03:28 (SD=00:20:33) hours, minutes and seconds for the expert dyads, 00:26:18 (SD=00:15:08) hours, minutes and seconds for the novice dyads and 00:49:57 (SD=00:20:57) hours, minutes and seconds for the individual searchers. A one way ANOVA revealed a main effect of Group ($F(2, 39) = 13.455$, $p < .001$, ges = .408). Novice dyads took significantly less time to search than the novice individuals

(p = .007) and the expert dyads (p < .001). No other contrasts were significant (p = .434).

Measures of Exhaustiveness

Probability of Examining ROIs

The probability of examining ROIs was computed for dyads after summing across each member of the dyad. The expert dyads performed at ceiling with every pair fixating every ROI from each ROI type (those requiring opening and/or moving object(s) and those requiring haptic examination). For this reason they were not included in the analysis of the probability of examining ROIs. For both the novice dyads and the novice individuals, the probability of examining either type of ROI, was significantly different from ceiling (*ps* < .001). A two (ROI type) by two (Group type) ANOVA revealed a main effect for ROI type ($F(1, 32) = 18.048$, $p < .001$, ges = .165 see Figure 4.3a) and Group ($F(1, 32) = 37.344$, $p < .001$, ges = .431). Participants were more likely to fixate ROIs where haptic examination was required than those requiring visual inspection after opening and/or moving objects. Novice dyads examined more ROIs than novice individuals. The interaction between Group and ROI type was also significant ($F(1, 32) = 11.385$, $p = .002$, ges = .111). A series of one-way ANOVAs revealed that Group was significant for both ROI types (*ps* < .001); but the difference between Groups was larger for ROIs requiring opening and/or moving object(s) than ROIs that required haptic examination. Furthermore, ROI type was significant for both Groups (*ps* < .014); novice dyads were more likely to examine ROIs requiring opening and/or moving object(s) than the novice individuals (see Figure 4.3a).

Search Completeness

For those ROIs that participants fixated, the likelihood that exhaustive search followed can be calculated as the mean deviation from that required for an exhaustive search of ROIs of that type. For example, if a participant performed only a visual inspection of an ROI that required haptic search then that would give a deviation score of -2, as they had failed to move and/or open objects AND conduct a haptic examination. A score was calculated for each participant/dyad by subtracting the score associated with the actual search action performed at each particular ROI from that required for an exhaustive search. A mean score was calculated across all ROIs of each type. A two (ROI type) by two (Group type) ANOVA revealed a main effect for ROI type ($F(1, 39) = 52.008$, $p < .001$, ges = .309, see Figure 4.3b) and Group ($F(2, 39) = 9.456$, $p < .001$, ges = .244). Participants were more likely to conduct an exhaustive search of ROIs requiring opening and/or moving objects(s) than haptic examination. The Expert dyads searched more exhaustively than the novice dyads ($p < .001$) but the differences between expert dyads and novice individuals, and novice dyads and novice individuals did not reach significance (*ps* > .068).

The interaction between ROI type and Group was also significant ($F(2, 39) = 3.276$, $p = .048$, ges = .053). The Groups differed for ROIs requiring haptic examination ($F(2, 39) = 11.321$, $p < .001$, ges = .367) but not for ROIs requiring opening/moving object(s) ($F(2, 39) = 2.466$, $p = .098$, ges = .112). There was an effect of exhaustiveness of search for ROI type for all Groups (*ps* < .014). However, the failure to complete haptic search was more apparent for novice dyads than for the expert dyads. The same contrast was marginally significant for novice individuals and novice dyads.

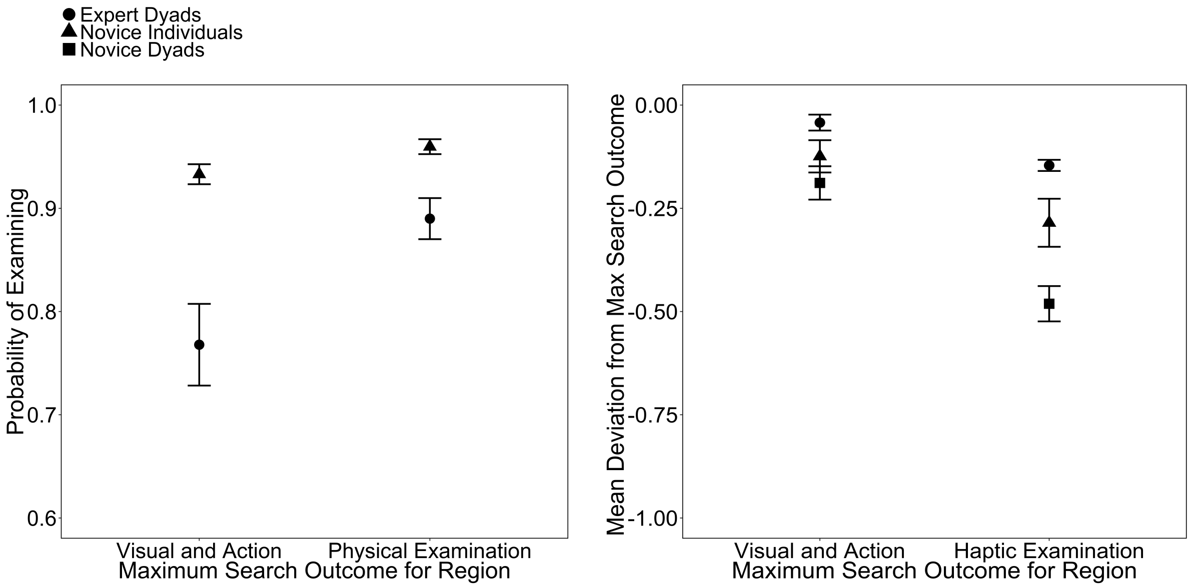

Figure 4.3 Measures of Exhaustiveness. Probability of examining ROIs that should be searched by visual inspection following opening and/or moving objects (action) or haptic examination for exhaustive search to be complete (with standard error of the mean; SEM) (a) and Mean deviation from the maximum possible search outcome for each ROI (with SEM) for novice individuals, novice dyads and expert dyads (b).

Additional Analysis

To provide a more complete understanding of RS behaviour, two further analyses are provided. First, an analysis of the extent to which dyads overlapped in their patterns of inspection. Second, an analysis of basic eye movement measures. Both analyses are computed irrespective of type of ROI, as determination of ROI requires at least some period of initial inspection in order for the determination to be made.

Overlap

An overlap of 0% would be consistent with pairs splitting the rummage task perfectly, whereas an overlap of 100% would indicate pairs searched the same ROIs. The novice and expert dyads differed in their amount of overlap ($t(12.44) = -3.03$, $p = .01$). Expert dyads overlapped in fixating 91% (SD=19%) of the same ROIs and the novice dyads overlapped in fixating 67% (SD=20%) of ROIs. Furthermore, by taking the probability of fixating ROI for the individual searchers (82% across both ROI types), the joint probability of two independent searchers both fixating 82% of ROI can be calculated. This gives us an estimate of 67% of overlap between two individuals conducting independent RS. Whereas expert dyads overlap a lot more in their RS as per their training, the performance of novice dyads is consistent with them performing two independent searches.

Eye Movement Measures

In addition to allowing us to measure exhaustiveness, the eye movement data also allow measurement of the information processing during RS. Eye movements for dyads were from individuals within dyads. Eye movement analyses were conducted using a one-way ANOVA repeated over the Group factor (Group: novice individuals, novice dyads, novice experts). Significant main effects were broken down using Bonferroni corrected t-tests. Three eye movement measures were examined based on gaze behaviour, i.e. the period of time spent fixating a ROI before moving the eyes away from that ROI; (1) Mean number of gazes, calculated as the average number of gazes made during the task (2) Mean gaze duration, calculated as the average length of time for a gaze and (3) Proportion of time spent fixating ROIs, calculated as the proportion of time spent fixating ROI relative to the total gaze time overall.

For mean number of gazes, the main effect of Group was not significant ($F(2, 71) = 1.355$, $p = .264$, ges = .037, see Figure 4.4a). For mean gaze duration, the main effect of Group was significant ($F(2, 71) = 40.295$, $p < .001$, ges = .531, see Figure 4.4b), with novice dyads making shorter gaze durations than both novice individuals ($p < .001$) and expert dyads ($p < .001$), however, there was no difference between the novice individuals and the expert dyads in relation to gaze (p = . 99). For the proportion of time spent fixating ROIs, the main effect of Group was significant ($F(2, 71) = 3.171$, $p = .048$, ges = .082 see Figure 4.4c), but no pairwise contrasts reached significance. Overall, these eye movement analyses indicate that although there were no differences in the proportion of time spent fixating ROIs compared to fixating non-ROIs, nor a difference in the number of gazes made between participant groups, the experts spent longer visually processing the ROIs than novice dyads. However this difference was not true for the experts and the novice individuals. To be clear, it appears that the experts and novices were equally likely to visually engage with a ROI, but when doing so, experts spent longer than novice dyads visually processing it.

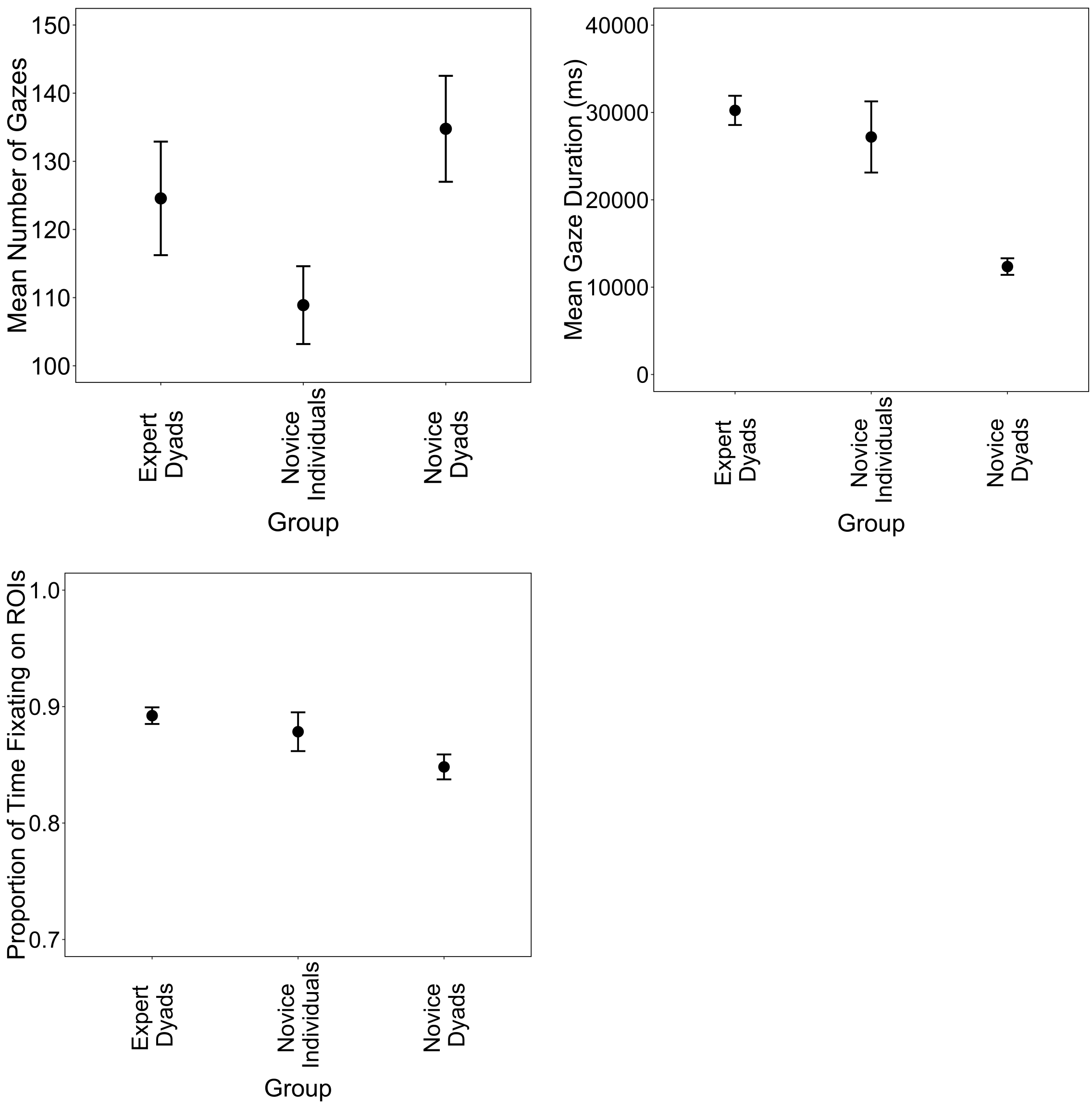

Figure 4.4 Eye movement measures. Proportion of time spent fixating ROIs (with SEM) (a), Mean gaze duration (with SEM) (b), and Mean number of gazes (with SEM) (c), for novice individuals, novice dyads and expert dyads.

Discussion

The goal of Chapter 4 was to explore RS when searching for clues to a crime. In particular, two reasons for failing to find targets in a RS task were explored: is it because ROIs holding targets are not attended, or is it because the actions conducted are insufficient to thoroughly search the ROIs that are attended. These questions were explored in relation to experts and novices working in dyads, justified on the grounds that expert rummage searchers typically work in pairs. They were also explored in relation to novice individuals searching alone.

First a framework was developed for classifying the pattern of fixations to ROIs and the motoric actions made to explore ROIs. The framework provides a conceptual distinction between three classes of action that, if poorly executed during RS, lead to targets being missed. At the outset, the presumption was that experts would show the upper limit of performance in RS, whereas novices would miss targets through failing to attend and explore ROIs exhaustively.

The data showed expert dyads were at ceiling for target detection, to search all possible ROIs and to be very close to ensuring all ROIs fixated were searched exhaustively. In addition, and consistent with their training, expert dyads overlapped on almost all the ROIs they searched, consistent with use of a double-checking procedure. With respect to the three attributes required for excellent RS that informed the framework used for coding performance, these data are consistent with experts 1) following a structured path, 2) searching all possible target locations along that path and 3) doing so exhaustively, therefore encoding the affordances of environments to hide objects. Note, although in Chapter 4 the physical route taken by searchers was not recorded (unlike in Chapter 3), the fact that the experts viewed all ROIs suggests that they were systematic in their path. If they hadn't been, they likely would have missed ROIs. It is interesting to note that when deciding on a search strategy, clear differences could be observed between the expert and novice dyads. The novice dyads typically opted to split the rooms so that each member of the dyad started their search in a different location, searching different regions of the space at the same time. The expert dyads, however, followed each other and remained close to each other throughout, as is also revealed in their tendency to double check each other. Although searching different regions of space at the same time, likewise to the independent dyads in Chapter 3, the novice dyads were still working as a dyad and compared to searching individually, would have been influenced by the other searcher in terms of where they started their search, their speed of search, and when they finished their search.

Evidence that expert dyads used an exhaustive RS strategy (and spent longer doing so too) is reminiscent of that reported in other security related searches. Biggs et al., (2013) and Biggs and Mitroff (2013) tested airport security luggage screeners and found that in comparison to student searchers, airport security searchers were more accurate, took less time to locate a target, and critically, spent longer searching the display before terminating their search (Biggs et al., 2013). A part of the skill of being an expert is to calibrate how much time is required to search effectively and to not terminate search before that time occurs.

With respect to the novice rummage searchers (searching in dyads or as individuals), it is important to determine the best novice group to compare with the expert dyads. The data suggest that working in dyads, relative to working alone, did not affect the number of targets found, the number of gazes made, the proportion of time fixating ROIs, or the level of exhaustiveness with ROIs. In contrast, it did reduce the time spent searching and the mean fixation durations, while increasing the number of ROIs examined. Although the novice dyads examined more ROIs than novice individuals, they become more superficial in their search. Where novice dyads and individuals differ with experts, the most conservative comparison tends to be with novice individuals.

Relative to experts, novices missed targets through failing to examine some ROIs, and failed to search those examined exhaustively. The failure to search ROIs exhaustively was more marked for those requiring haptic examination confirming the prediction that as the motoric effort required for searching ROIs increased, participants would be less likely to be exhaustive. In showing systematic differences between experts and novices during RS, it is concluded that the framework used for coding performance has been helpful in determining sources of error in novice RS that lead to missing targets.

It seems likely that the experts just 'see' the search environment differently to novices. As such, knowing what to search and appreciating the affordances of how environments can hide targets are skills improved by training. Visual expertise is often considered as being reflected in the familiarity and exposure to classes of stimuli [expertise effects have previously been shown to exist in a range of domains such as medicine (Krupinski et al., 1996, 2006; Kundel & Nodine, 1983; Kundel et al., 2007; Kundel & La Follette, 1972; Nodine et al., 2002), sport (Howard, Troscianko, & Gilchrist, 2010; Land & McLeod, 2000) and reading music (Rayner & Pollatsek, 1997)]. Furthermore, the present study has shown effects of experience on decision-making with respect to the presence of IEDs, consistent with learning the affordances of environments to hide targets (Godwin et al., 2015). With respect to the current study, expertise is manifested in understanding of the affordance of a residential house to hide targets. To some extent, this is a surprising outcome given that a residential house was a familiar environment to all participants. It cannot, therefore, be the visual context itself with which we need to become more familiar but its potential to afford hiding. Evidently, it is the act of exhaustive searching within this context that requires training.

The present experiment was novel. The reality of RS was maintained in the real world, at the expense of close control over that environment, to explore a framework for understanding rummage searching. Testing performance using an analytic framework for RS revealed data consistent with effects of expertise in RS. It is concluded that the framework for studying RS provided here has been effective in revealing a difference between novice and experts: experts find more targets than novices because they examine more ROIs and examine ROIs more exhaustively. The failure to fixate ROIs might occur through (1) failing to determine and follow a path that brings all ROIs within a search path, or (2) failing to coordinate search within the ESW of the search path. It is for future work to discriminate between these alternatives.

Finally, in outlining the framework for understanding RS and showing how it is helpful in discriminating expert from novice searchers, there are two important challenges for future work. First, does the framework generalise such that it is similarly helpful in explaining data from other RS scenarios? Second, is the framework helpful in differentiating separate classes of action that need to be considered separately in training?

Chapter 5

General Discussion

The motivation of this thesis was to explore search behaviour in complex real world tasks. In particular, the real-world search tasks of airport X-ray security screening on a computer display, searching for clues to a crime or IEDs in patrol on a large area of grassland, and rummage searching both visually and haptically. In each case, experimental variants of these complex real-world tasks were explored, as well as how searching alone or with a partner, changed search behaviour.

Chapter 1

Searching in the real world *can* be a simple task, but more often it is a complex and challenging one. Chapter 1 considered the basic models of visual search, and how viewing photographs and scenes added context to this. Furthermore, a number of factors were discussed as having an influence over efficiency of search in real world tasks, from differences in the types and quantity of targets sought, to the biases and decisions associated with more complex search tasks. These were reviewed in Chapter 1, in the context of trying to understand individuals searching for targets alone and when searching with others.

Chapter 2

In terms of the complex search task of airport security screening, Chapter 1 revealed a number of factors to be important. Of particular interest to Chapter 2 was that search can be incomplete when searching for infrequently occurring targets, i.e. a rare IED (the prevalence effect; Godwin et al., 2010; 2015; Wolfe et al., 2007; 2005) and when searching for two target types simultaneously, i.e. guns and IEDs (the dual-target cost; Menneer et al., 2007).

Chapter 2 took these two factors and explored them in relation to how working in a pair might influence search in a task similar to airport security screening. Although modelled on the real world task, the task examined is an analogue of the real world search task, it is important to note that although encompassing many aspects of the airport security screening task, the experiments in Chapter 2 did not use x-ray images and the stimuli did not overlap. The experimental set up acted as a simplified version of the task. Therefore, although considered here in relation to airport security screening, it is also important to note that the implications of Chapter 2 can also be considered in relation to other real world complex search tasks, such as searching medical images for abnormalities, and other more basic search theory too. Chapter 2 is important for its applied implications, but also in a theoretical sense, as it builds on basic visual search findings.

Given that searching for two target types simultaneously has been found to lead to a dual-target cost in performance, suggestions have been made to encourage specialist screeners to search for specific categories of threat only, e.g. one searches for guns and another searches for IEDs (Menneer et al., 2007; 2009; 2012) Chapter 2 considered the implications of more than one searcher working along side each other, and how *expectations* influence search. To establish a baseline, in the first experiment individual searchers were tested, in a task which explored the expectations that develop when participants search for targets of differing frequencies (i.e., the prevalence effect). As discussed in Chapter 1, it has been suggested that the prevalence effect occurs due to a response bias (Chun & Wolfe, 1996; Godwin et al., 2010, 2015; Schwark et al., 2013; Wolfe & Van Wert, 2010), and so Chapter 2 also explored the source of this bias. The question considered was could the prevalence effect be alleviated by either bottom-up priming in the form of target repetition, or top-down expectations regarding the likelihood of targets occurring. This was tested in a standard visual search task of find the 'T' shape amongst 'L' shapes. Prevalence levels alternated between two 'slides' of a trial (allowing participants to build an expectation for each slide) and colour was used to distinguish two conditions of bottom-up priming, 1) the same-colour condition, where participants received bottom up priming for the low prevalence target through repetition of the high prevalence target, and 2) the alternating-colour condition, where the prevalence effect was expected to show.

In the second experiment of Chapter 2, participants were tested alongside a simulated 'partner' programmed to act with a response bias. This 'partner' was either susceptible to making more misses or making more false alarms. This experiment took a different approach to exploring expectations, as shifts that arise based upon the basis of a 'partner's' performance could only be a top-down influence. This was also tested in a find the 'T' amongst 'L's task, with the participant told that their 'partner' had already searched the display for their coloured 'T' shape, and that they were to search for a different coloured 'T' shape.

Key Findings

Chapter 2 revealed four key findings. 1) The Experiments in Chapter 2 both revealed the typical differences expected to be found between target present and target absent trials. This meant that the task was suitable to assess other visual search behaviour. 2) Experiment 1 found clear evidence for bottom-up priming to influence the prevalence effect in an individual visual search task. The prevalence effect could be modulated through repetition of the target. 3) Experiment 1 found weak evidence of a top down influence, only through a null effect of a colour interaction in one eye movement measure. Finally, 4) Experiment 2 found one marginal effect of response criterion, with participant fixation durations increasing as their 'partner' made more misses. This suggested that observing misses may not have altered the outcome of behaviour, but began to create uncertainty in how the participant conducted their search.

Implications and Future Study

The Experiments in Chapter 2 provide reason to pursue these issues further. The influence of bottom-up priming is clear, but the influence of top-down expectations is a little less clear. Future work should tease apart the influence of bottom-up priming and top-down expectations in an individual search task, while increasing the credibility of the 'partner' in the paired search task may lead to more concrete conclusions.

In terms of the implications to the applied search task of X-ray airport security screening, further research is needed before implementing specialist screeners. Furthermore, others (e.g. Mitroff et al., 2015) have recently made suggestions to move the airport screening process to a remote location, to help prevent another issue important to airport security screening, searching for multiple targets in an image. The idea behind this suggestion being that in a remote screening scenario, bags can be viewed by more than one screener. The findings presented in Chapter 2 suggest further experimentation is required before this is implemented either (although Mitroff et al. (2015) also notes this to be true), as searching for multiple targets together is likely to be subject to the same influence as searching for different target types together; searchers may change their own search behaviour based on the search behaviour of their colleague. Furthermore, airports in the UK are currently advancing and introducing ‘Centralised Image Projection’, whereby a second searcher checks the performance of an initial searcher. The findings of Chapter 2 have important implications for this too, as it may be that the second searcher checking performance may be biased by the performance of the initial searcher, rather than performing based on the actual display. Finally, at this stage what can be suggested is that if working together, encouraging participants to actively ignore another’s target and the quality of their search will be important to ensure efficient search yourself. If further experimentation reveals a stronger effect of a partner’s response bias, then this could a positive outcome, if it encourages the searcher to become more conservative, but will be a negative outcome if it encourages the searcher to become more liberal instead.

Chapter 3

Chapter 3 considered the type of complex real world search task where police search teams comb open ground for clues to a crime, or soldiers patrol high-risk routes searching for IEDs on a large area of grassland. In addition to the issues explored in Chapter 2, this form of search task is also of a larger-scale, requiring eye, head and body movements (Foulsham et al., 2014; Gilchrist et al., 2001; Jiang et al., Won et al., 2014), integrates aspects of foraging and searching for an unknown number of targets (Cain et al., 2012; Wolfe, 2013) and requires defining and following a search path (Koopman, 1946, 1980). Furthermore, the implication of paired searching is also of importance. Chapter 3 (and Chapter 4 also) moved beyond standard visual search tasks and required the development of a novel analytic framework, before experimentation could take place. Chapter 3 conceptualised a framework in which to explore the search paths, developed in the context of theories of search and rescue (Koopman, 1946, 1980) and used Fourier analysis to create a number of measures in the frequency domain to represent movement, which could then be considered in relation to accuracy and search time.

Using this novel analytic framework, Chapter 3 explored what leads to good performance in this type of task, and what search strategies maximise the number of targets detected. To do this, participants were required to search for an unknown number of coin targets within a bounded area of grassland. The participant's objective was to find all the coins, and the task instructions did not define any search strategy, only that they should terminate search when satisfied that search was complete. Furthermore, in the same experimental set up, Chapter 3 also explored what happens when participants are subjected to variations in target density/prevalence, and what happens when two searchers are required to search together. As in Chapter 2, the experimental set up for Chapter 3 acts as an analogue for the real world task. Chapter 3 does, however, go further than Chapter 2 in mimicking the real world task, but does still have its limitations (such as the use of the well-defined rectangular grid, and the type of targets used, as discussed under implications).

Key Findings

Chapter 3 revealed seven key findings. 1) The analytic framework was successful in capturing search paths and assessing performance beyond mere accuracy and time. 2) The task was found to be challenging, however targets were detected at a fixed rate, suggesting it was not subject to inappropriate satisfaction of search (Tuddenham, 1962)/subsequent search misses (Adamo et al., 2013), or reduced vigilance. 3) Participants tended to use an 'S' shaped search pattern with a common width of search lane. This strategy minimized memory load and maximized the possibility of an exhaustive search. 4) Increased accuracy was associated with slower (i.e. longer time before changing direction), but also more variable (i.e. stopping to check targets), search speed, though only when participants moved along the length (as opposed to across the width) of the search area. This allowed for eye, head and body movements to be made across the search width. 5) Participants adopted this 'S' shaped search pattern across variations in target prevalence/density, and when there was more than one searcher. 6) Furthermore, accuracy, total search time and rate of detection were also unchanged by target prevalence/density. Finally 7) While dyads were more accurate than individuals, dyads that opted to conduct two independent searches were more accurate than those who opted to split the search space. This is a critical point. Integration of search effort comes at a cost, and although the underlying source of this cost is currently unknown, the best performance clearly forms from the benefit of independent probability summation of two searchers.

Implications and Future Study

The Experiments in Chapter 3 revealed a number of important findings, however, they also raised further questions, which must be examined in future work. First, the $75m^2$ grid was larger than typical visual search studies, however, the search area was rectangular and sufficiently small that the cones used to define the search boundary could have acted as landmarks to guide searching. Future studies could explore how strategy is influenced by the shape and size of the search area. Furthermore, the visibility of the cones acting as a boundary could be altered to explore the influence of external markers on search behaviour.

Second, it would be interesting to investigate whether changes in the grid influence the performance of the dyads. Are they still able to outperform individuals when the search space is not rectangular or clearly defined? This may lead to further costs in coordination, and may lead to a higher proportion of dyads opting to search 'independently' of each other. Also of interest is the decision making process of the dyads, and what happens when a dyad increases to a team, as is often the case in the real world task. Finally, what happens to search strategy and performance when dyads are given explicit instructions regarding an adopted search strategy, or what happens when they are given a time pressure in which to search within.

A third area warranting investigation is the discriminability of the targets. What happens when the targets change in size and/or degree of concealment? What's more, the prevalence manipulation used in the present thesis gave an initial insight, but further testing is needed here too. In Chapter 3, all participants took part in each prevalence condition, further experimentation could use a blocked design and subject participants to multiple search grids of the same prevalence level. Further testing is needed before we can be certain about the impact of the prevalence effect in this form of search task. What is clear, however, is in the current experimental set up, participants' search path is determined independent of target prevalence.

Chapter 3 demonstrated the use of a novel analytic tool to capture movement performance in a large-scale search task. This could be built upon in future study by integrating eye movements also. In terms of the implications for the real world search task, the present study has revealed some intriguing results. However, the present study was not carried out under safety critical conditions and therefore participants did not experience the heightened emotions or gravity of the consequences of missing targets that often coincide with the real world version of the task. Building on from Chapter 2, which suggested that searchers should be encouraged to actively ignore another searcher's target and quality of search, Chapter 3 suggests that it might be helpful to encourage participants who are physically working together, to work independently of each other, to fully overcome any cost in dyadic coordination.

Chapter 4

Chapter 4 considered the complex real world search task of Rummage Search. We all engage in forms of RS in everyday life, such as looking for our keys, however, in security critical scenarios, RS is conducted to detect threatening targets which are purposely hidden. The police, military and border agencies perform RS to find weapons, drugs and IEDs in locations such as houses, cars, and venues. RS builds on the previous tasks in the present thesis with the addition of physical motoric search of objects which may be hidden or partially hidden, may overlap or may be broken down into parts. As with previous chapters, the implication of paired searching is also of importance, as expert searchers often carry out RS in pairs within larger search teams. As in Chapter 3, Chapter 4 moved beyond standard visual search tasks and required a novel analytic framework to be developed, before experimentation could take place. Chapter 4 conceptualised a framework integrating eye movements in order to code and examine physical action. A number of measures were created to examine the 'exhaustiveness' of search, in addition to typical eye movement and behavioural measures.

Chapter 4 developed and explored RS within a framework based on level of motoric effort; 1) inspection of visible areas (i.e., a table top); 2), visual inspection that would be improved by moving or opening objects (i.e. a cupboard door) and 3) and third, haptic examination of objects (e.g., the compressing of a cushion) where visual inspection alone would have been incomplete. This was used to measure 'exhaustiveness' in participants and was explored within a group of expert military searchers (who typically work in dyads), in addition to novice individuals and novice dyads. Participants engaged in RS of a residential house for incriminating evidence relating to a crime. Participants were given a briefing and asked to search for targets relevant to the brief. They were instructed to locate any firearms or money associated with a robbery at a bank, and any drugs that may be present at the property too. Participants did not know the number, location or distribution of potential targets. As in previous Chapters, the experimental set up for Chapter 4 acted as an analogue for the real world task. Chapter 4 does, however, go further than both Chapter 2 and Chapter 3 in mimicking the real world task. It is the closest it comes to a simple simulation of reality, however the threats were not real and the house was designed for certain challenges.

Key Findings

Chapter 4 revealed six key findings. 1) The analytic framework provided a conceptual distinction between three classes of action that, if poorly executed during RS, led to targets being missed. 2) Expert dyads were very good at the task, conducting a slowed, double-checking exhaustive search strategy. 3) Experts demonstrated three key attributes required for efficient RS; Following a structured path, searching all possible target locations along that path and doing so exhaustively, therefore encoding the affordances of environments to hide objects. 4) Relative to the experts, novice participants missed targets through both failing to examine ROIs, and failing to search those examined exhaustively. 5) This failure to search ROIs exhaustively was more marked for ROI requiring haptic examination, suggesting that as more motoric effort is required, participants are less likely to be exhaustive. Finally 6) Comparing novice individuals to novice dyads, both groups performed similarly across the majority of measures examined. However, dyads inspected more ROIs, but spent less time searching and made shorter fixations, suggesting although they examine more than the novice individuals, the novice dyads become more superficial in their search (i.e. less exhaustive).

Implications and Future Study

Chapter 4 revealed that a framework based on motoric effort was able to conceptualise RS. However, whether this framework generalizes to other RS scenarios (such as a car or venue) is currently unknown. Future work could start to explore this further using a single room, such as an office, in order to control the complexity of the environment in terms of colour, form and position of specific items. Furthermore for efficient RS, searchers should follow a structured path. It is important to note, however, that although the physical route taken by searchers was not recorded (unlike in Chapter 3), the fact that the experts viewed all ROIs suggested that they were systematic in their path. However, future work could aim to integrate the use of a tracking system in RS. The method used in Chapter 3 would not be suitable for the RS scenario (due to the size of the staff and the means by how coordinates are calculated) but perhaps another device could be suitably implemented into the task.

Chapter 4 also revealed that training is important in RS. It could be argued that everyone is familiar with a house scenario, but only the experts were able to efficiently search it. A further question would be does the failure in novices to fixate ROIs occur through either failing to determine and follow a path that brings all ROI within a search path or failing to coordinate search within the search path. Future work should discriminate this difference.

As in Chapter 3, questions remain surrounding target prevalence, and whether a low prevalence of targets influences the degree of exhaustiveness in RS, in both visual and haptic search. Chapter 4 already addressed the issue of concealment, but further experimentation should explore the issue of size, and what happens if a target is very small (i.e. a sim card), or smaller than expected to be. In terms of the implications for the real world search task, it would be interesting to explore whether the framework outlined is helpful in training those who conduct professional RS searches.

Conclusions

Two main conclusions are made from the experiments in this thesis. First, in regards to searching in dyads, Chapter 2 suggested that searchers should be encouraged to actively ignore another searcher's target and quality of search, while Chapter 3 suggested pairs should be encouraged to work independently of each other. Adding to this, Chapter 4 suggests this may be true in novice searchers, but that the key to coordination in dyads may be training. There is evidently a cost to working with someone else in a search task, and it is not clear that there is a substantial benefit to working as a dyad. Chapter 3 found an example of dyads outperforming individuals, but even then the dyads were at their best when they searched independently of each other. When they attempt to coordinate they incur a cost, which could become more evident in a less structured search area. Chapter 4 found that the dyads examined more than the individuals, but were less exhaustive with what they did examine. The benefit of paired searching was tempered by the fact that the dyads became more superficial in their search behaviour and did not search as exhaustively as the independent searchers, who behaved more similarly to the expert searchers. Furthermore, the expert searchers in Chapter 4 appear to be adopting a form of *coordinated independent search*, in that they are trained to check each other and overlap to search all possible locations, rather than attempting to split the search space spatially. It is important to note, however, that for Chapters 3 and 4, although elicting to either split the search space spatially, to overlap each other, or to search independently of each other, dyads were still part of a dyad and even if searching 'independently', would have been subject to influence in terms of where where they started their search, their speed of search, and when they finished their search. In other words, the independent dyads were never truly independent.

Second, the present thesis has demonstrated two examples of novel analytic frameworks used to explore behaviour in two different forms of search task. This was necessary as a first step to addressing the problems, before any data could even be collected. In this sense, the experiments run in this thesis are the first of many which could utilize this technique. Future work might take aspects of each of the solutions demonstrated in Chapters 3 and 4, so that behavioural measures could be integrated with eye movements, physical action *and* search path analysis. This could deliver a rich dataset for searching in complex environments, and form the basis of a theory of searching for targets in the real world. Future studies should also test the generalizability of this framework. In typical visual search studies (as in Chapter 2), it is relatively straightforward to run multiple trials and collect a wealth of data. In more demanding experiments (such as in Chapters 3 and 4) a lot of time and effort is required to both collect the data and to code it, making multiple 'trials' more of a challenge. It remains possible the data might change with more complicated arrays, and a question of real interest for the future is how does the data change with changes in target shape, density, discriminability etc. The frameworks outlined in this thesis can be integrated and utilized to help explore this question.

Appendices

Appendix A. The ROIs used in Chapter 4

ROIs listed in clockwise direction starting from the door of each room. Target locations also indicated. Note, in the Kitchen the ROIs cabinet, drawer 1, drawer 2 and cupboard are part of a larger item of furniture and the work surfaces have been split into three. All other ROI are separate items.

Room	ROI	Search output required	Target
Kitchen	Vacuum Cleaner	Visual inspection, including following opening of parts and moving whole vacuum cleaner.	
	Fridge Freezer	Visual inspection, including following opening of door and moving of water bottles.	
	Worksurface 1	Visual inspection, including following moving of items on the work surface.	
	Cupboard 1	Visual inspection, including following opening of door and moving of crockery AND haptic examination of hard to see space.	
	Cupboard 2	Visual inspection, including following opening of door and moving of crockery AND haptic examination of hard to see space.	
	Cupboard 3	Visual inspection, including following opening of door and moving of crockery AND haptic examination of hard to see space.	
	Bag	Visual inspection, including following opening of bag and moving contents (e.g. paper towels) and whole bag AND haptic examination (e.g. sides of bag and material inside).	
	Window Sill	Visual inspection, including following moving of objects on the windowsill.	
	Worksurface 2	Visual inspection, including following moving of items on the work surface.	
	Mini Fridge	Visual inspection, including following opening of door.	
	Drawer 1	Visual inspection, including following opening of door and moving of objects (e.g. cutlery) AND haptic examination of objects (e.g. packet of antibacterial wipes).	
	Drawer 2	Visual inspection, including following opening of door.	
	Cupboard 4	Visual inspection, including following opening of door and moving of objects (e.g. crockery) AND haptic examination of hard to see space.	
	Cupboard 5	Visual inspection, including following opening of door and moving of objects (e.g. crockery) AND haptic examination of hard to see space.	
	Worksurface 3	Visual inspection, including following moving of items on the work surface.	
	Drawer 3	Visual inspection, including following opening of door and moving of objects (e.g. cutlery) AND haptic examination (e.g. packet of antibacterial wipes).	
	Cupboard 6	Visual inspection, including following opening of door and moving of objects (e.g. crockery) AND haptic examination of hard to see space.	

Room	ROI	Search output required	Target
	Bin	Visual inspection, including following moving of rubbish and whole bin.	
	Drawer 4	Visual inspection, including following opening of door and moving of objects (e.g. cutlery) AND haptic examination of objects (e.g. packet of antibacterial wipes).	
	Boiler	Visual inspection following opening of door.	
	Cupboard 7	Visual inspection, including following opening of door and moving of objects (e.g. crockery) AND haptic examination of hard to see space.	
	Clock	Visual inspection following moving of whole clock.	
	Radiator	Visual inspection AND haptic examination of space behind radiator.	
Living Room	Magazine Rack	Visual inspection, including following moving of individual magazines and whole magazine rack.	
	Cabinet	Visual inspection following opening of door and moving ornaments.	
	Drawer 1	Visual inspection, including following opening of drawer.	
	Drawer 2	Visual inspection, including following opening of drawer.	
	Cupboard	Visual inspection, including following opening of door and moving of vinyl records AND haptic examination of hard to see space.	Large pack of drugs inside a box of vinyl records inside cupboard
	Flower Pot	Visual inspection, including following moving of flowers and whole flowerpot.	
	Picture	Visual inspection, including following moving of whole picture.	
	Fireplace	Visual inspection, including following moving of items on the fireplace.	
	TV Unit	Visual inspection, including following moving of items on the unit.	
	Rug	Visual inspection, including following moving of whole rug.	
	Lamp	Visual inspection, including following moving of whole lamp AND haptic examination of inside of lampshade.	
	Bin	Visual inspection, including following moving of items inside bin or whole bin.	
	Window sill	Visual inspection, including following moving of objects on the windowsill.	
	Radiator	Visual inspection AND haptic examination of space behind radiator.	
	Bookcase	Visual inspection, including following moving of books AND haptic examination of hard to see spaces.	Small roll of money inside one of the game boxes on the bookshelf

Room	ROI	Search output required	Target
Large Bedroom	Glass Table	Visual inspection, including following moving of items on the table.	
	Sofa	Visual inspection, including following moving of cushions AND haptic examination of cushions.	
	Chair	Visual inspection, including following moving of cushions AND haptic examination of cushions.	
	Wardrobe	Visual inspection, including following opening of door and moving of objects (e.g. clothes and bedding) AND haptic examination of objects (e.g. clothes and bedding).	
	Chest of Drawers	Visual inspection, including following opening of drawers and moving of clothes AND haptic examination of clothes.	
	TV	Visual inspection following moving of whole TV.	
	Bedside Table	Visual inspection, including following opening of door and moving of boxes AND haptic examination of the underside of table.	
	Window sill	Visual inspection following moving of items on windowsill.	
	Radiator	Visual inspection and haptic examination of space behind radiator.	
	Bed	Visual inspection, including following moving of objects (e.g. pillows) AND haptic examination of bedding and underside of bed.	
Small Bedroom	Bag	Visual inspection, including following opening of bag and moving (e.g. of sheets) AND haptic examination (e.g. sides of bag and sheets inside).	A long barreled weapon (rifle) inside a bag under the bed
	Desk	Visual inspection, including following opening of drawers and moving objects (e.g. stationary) AND haptic examination of hard to see spaces.	
	Bean Bag	Visual inspection, including following opening of outer cover and moving whole beanbag AND haptic examination the beanbag.	A large bundle of money inside the beanbag
	Window sill	Visual inspection, including following moving of items on windowsill.	A small package of drugs placed on the windowsill
	Radiator	Visual inspection AND haptic examination of space behind radiator.	
	Chair	Visual inspection, including following moving of cushions AND haptic examination of cushions.	
	Guitar	Visual inspection, including following moving of whole guitar.	
	Cupboard	Visual inspection, including following opening of door or moving of toys AND haptic examination of hard to see spaces.	A short barreled weapon (pistol) inside the cupboard
	Wardrobe	Visual inspection, including following opening of door or moving of clothes and bedding AND haptic examination of clothes and bedding.	
	Bed	Visual inspection, including following moving of objects (e.g. of pillows) AND haptic examination of bedding and underside of bed.	

References

Adamo, S. H., Cain, M. S., & Mitroff, S. R. (2013). Self-Induced Attentional Blink: A Cause of Errors in Multiple-Target Search. *Psychological Science*, *24*(October), 2569–2574. http://doi.org/10.1177/0956797613497970

Adamo, S. H., Cain, M. S., & Mitroff, S. R. (2015). Targets Need Their Own Personal Space: Effects of Clutter on Multiple-Target Search Accuracy. *Perception*, *0*(0), 1–12. http://doi.org/10.1177/0301006615594921

Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous Control of Attention by Multiple Working Memory Representations. *Psychological Science*, *23*(8), 887–898. http://doi.org/10.1177/0956797612439068

Berbaum, K., Franken, E., Dorfman, D., & Caldwell, R. (2000). Role of faulty decision making in the satisfaction of search effect in chest radiography. *Academic Radiology*, *7*(12), 1098–1106. Retrieved from http://www.sciencedirect.com/science/article/pii/S107663320080063X

Berbaum, K., Franken, E., Dorfman, D., Rooholamini, S., Kathol, M., Barloon, T., … Montgomery, W. (1990). Satisfaction of search in diagnostic radiology. *Investigative Radiology*, *25*(2), 133–140. Retrieved from http://journals.lww.com/investigativeradiology/Abstract/1990/02000/Satisfaction_of_Search_in_Diagnostic_Radiology.6.aspx

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

Biederman, I. (1972). Perceiving real-world scenes. *Science*, *177*(43), 77–80. Retrieved from http://graphics.cs.cmu.edu/courses/P2P/PDF/Biederman_Science_1972.pdf

Biederman, I., Mezzanotte, R., & Rabinowitz, J. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*(2), 143–177. Retrieved from http://www.sciencedirect.com/science/article/pii/001002858290007X

Biggs, A., Adamo, S., Dowd, E., & Mitroff, S. (2015). Examining perceptual and conceptual set biases in multiple-target visual search. *Attention, Perception, & Psychophysics*, *77*(3), 844–855. Retrieved from http://link.springer.com/article/10.3758/s13414-014-0822-0

Biggs, A. T., Cain, M. S., Clark, K., Darling, E. F., & Mitroff, S. R. (2013). Assessing visual search performance differences between Transportation Security Administration Officers and nonprofessional visual searchers. *Visual Cognition*, *21*(3), 330–352. http://doi.org/10.1080/13506285.2013.790329

Biggs, A. T., & Mitroff, S. R. (2013). Different predictors of multiple-target search accuracy between nonprofessional and professional visual searchers. *Quarterly Journal of Experimental Psychology*, *67*(7), 1335–1348. http://doi.org/10.1080/17470218.2013.859715

Birkelo, C. C., Chamberlain, W. E., Phelps, P. S., Schools, P. E., Zacks, D., & Yerushalmy, J. (1947). Tuberculosis case finding: A comparison of the effectiveness of various roentgenographic and photofluorographic methods. *Journal of the American Medical Association*, *133*(6), 359–366. Retrieved from http://jama.jamanetwork.com/article.aspx?articleid=291546

Bono, J. E., Boles, T. L., Judge, T. a, & Lauver, K. J. (2002). The role of personality in task and relationship conflict. *Journal of Personality*, *70*(3), 311–344. http://doi.org/10.1111/1467-6494.05007

Brennan, A. A., & Enns, J. T. (2015). When two heads are better than one: Interactive versus independent benefits of collaborative cognition. *Psychonomic Bulletin & Review*, *22*(4), 1076–1082. http://doi.org/10.3758/s13423-014-0765-4

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, *106*(3), 1465–1477. http://doi.org/10.1016/j.cognition.2007.05.012

Buswell, G. (1935). *How people look at pictures*. Chicago: University of Chicago Press.

Cain, M. S., Adamo, S. H., & Mitroff, S. R. (2013). A taxonomy of errors in multiple-target visual search. *Visual Cognition*, *21*(7), 899–921. http://doi.org/10.1080/13506285.2013.843627

Cain, M. S., Biggs, A. T., Darling, E. F., & Mitroff, S. R. (2014). A little bit of history repeating: Splitting up multiple-target visual searches decreases second-target miss errors. *Journal of Experimental Psychology: Applied*, *20*(2), 112–125. http://doi.org/10.1037/xap0000014

Cain, M. S., Dunsmoor, J. E., LaBar, K. S., & Mitroff, S. R. (2011). Anticipatory anxiety hinders detection of a second target in dual-target search. *Psychological Science*, *22*(7), 866–71. http://doi.org/10.1177/0956797611412393

Cain, M. S., & Mitroff, S. R. (2013). Memory for found targets interferes with subsequent performance in multiple-target visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1398–1408. http://doi.org/10.1037/a0030726

Cain, M. S., Vul, E., Clark, K., & Mitroff, S. R. (2012). A bayesian optimal foraging model of human visual search. *Psychological Science*, *23*(9), 1047–54. http://doi.org/10.1177/0956797612440460

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, *57*(8), 1241–1261. http://doi.org/10.3758/BF03208380

Carrasco, M., & Frieder, K. S. (1997). Cortical magnification neutralized the eccentricity effect in visual search. *Vision Research*, *37*(1), 63–82. http://doi.org/10.1016/S0042-6989(96)00102-2

Carrasco, M., McLean, T. L., Katz, S. M., & Frieder, K. S. (1998). Feature asymmetries in visual search: Effects of display duration, target eccentricity, orientation and spatial frequency. *Vision Research*, *38*(3), 347–374. http://doi.org/10.1016/S0042-6989(97)00152-1

Castelhano, M. S., & Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(4), 753–763. http://doi.org/10.1037/0096-1523.33.4.753

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2016). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(3), 6. http://doi.org/10.1167/9.3.6.Introduction

Cave, K. R. (1999). The FeatureGate Model of Visual Selection. *Psychological Research*, *62*(2–3), 182–194. http://doi.org/10.1016/B978-012375731-9/50094-X

Chan, L. K. H., & Hayward, W. G. (2013). Visual search. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(4), 415–429. http://doi.org/10.1002/wcs.1235

Charnov, E. (1976). Optimal Foraging, the Marginal Value Theorem. *Theoretical Population Biology*, *9*(2), 129–136. Retrieved from http://www.sciencedirect.com/science/article/pii/004058097690040X

Chen, X. (2007). *Collaborative Visual Search*. Stony Brook University. Retrieved from Retrieved from SUNY Digital Repository. (Accession Order No. 1951/44780/000000160)

Chun, M. M., & Wolfe, J. M. (1996). Just say no: how are visual searches terminated when there is no target present? *Cognitive Psychology*, *30*(1), 39–78. http://doi.org/10.1006/cogp.1996.0002

Clark, K., Cain, M. S., Adcock, R. A., & Mitroff, S. R. (2014). Context matters: The structure of task goals affects accuracy in multiple-target visual search. *Applied Ergonomics*, *45*(3), 528–533. http://doi.org/10.1016/j.apergo.2013.07.008

Cohen, E., & Ruppin, E. (1999). From parallel to serial processing: A computational study of visual search. *Perception & Psychophysics*, *61*(7), 1449–1461. Retrieved from http://link.springer.com/article/10.3758/BF03206193

Dalmaijer, E., Van der Stigchel, S., Nijboer, T., Cornelissen, T., & Husain, M. (2015). CancellationTools: All-in-one software for administration and analysis of cancellation tasks. *Behavior Research*, *47*(4), 1065–1075. http://doi.org/10.3758/s13428-014-0522-7

Donnelly, N., Guest, R., Fairhurst, M., Potter, J., Deighton, A., & Patel, M. (1999). Developing algorithms to enhance the sensitivity of cancellation tests of visuospatial neglect. *Behavior Research Methods, Instruments, & Computers*, *31*(4), 668–673. http://doi.org/10.3758/BF03200743

Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433. Retrieved from http://psycnet.apa.org/journals/rev/96/3/433/

Eckstein, M. P. (2011). Visual search: a retrospective. *Journal of Vision*, *11*(5:14), 1–36. http://doi.org/10.1167/11.5.14

Ehinger, K., & Wolfe, J. (2016). When is it time to move to the next map? Optimal foraging in guided visual search. *Attention, Perception, & Psychophysics*, 1–17. Retrieved from http://link.springer.com/article/10.3758/s13414-016-1128-1

ESRI. (2011). ArcGIS Desktop: Release 10. Redlands, California: Environmental Systems Research Institute.

Ferreira, F., Apel, J., & Henderson, J. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, *12*(11), 405–410. http://doi.org/10.1016/j.tics.2008.07.007

Findlay, J., & Gilchrist, I. (2003). *Active vision: The psychology of looking and seeing*. Oxford: Oxford University Press.

Findlay, J. M. (1997). Saccade Target Selection During Visual-Search. *Vision Research*, *37*(5), 617–631. Retrieved from http://www.sciencedirect.com/science/article/pii/S0042698996002180

Findlay, J. M., Brown, V., & Gilchrist, I. D. (2001). Saccade target selection in visual search: The effect of information from the previous fixation. *Vision Research*, *41*(1), 87–95. http://doi.org/10.1016/S0042-6989(00)00236-4

Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science*, *18*(11), 943–947. http://doi.org/10.1111/j.1467-9280.2007.02006.x

Fleck, M. S., Samei, E., & Mitroff, S. R. (2010). Generalized "satisfaction of search": adverse influences on dual-target search accuracy. *Journal of Experimental Psychology. Applied*, *16*(1), 60–71. http://doi.org/10.1037/a0018629

Fougnie, D., Cormiea, S. M., Zhang, J., Alvarez, G., & Wolfe, J. (2015). Winter is coming: How humans forage in a temporally structured environment. *Journal of Vision*, *15*(11), 1–1. http://doi.org/10.1167/14.10.913

Foulsham, T., Chapman, C., Nasiopoulos, E., & Kingstone, A. (2014). Top-Down and Bottom-Up Aspects of Active Search in a Real-World Environment. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale*, *68*(1), 8–19. http://doi.org/10.1037/cep0000004

Foulsham, T., & Underwood, G. (2011). If Visual Saliency Predicts Search, Then Why? Evidence from Normal and Gaze-Contingent Search Tasks in Natural Scenes. *Cognitive Computation*, *3*(1), 48–63. http://doi.org/10.1007/s12559-010-9069-9

Gilchrist, I. D., & Harvey, M. (2006). Evidence for a systematic component within scan paths in visual search. *Visual Cognition*, *14*(4–8), 704–715. http://doi.org/10.1080/13506280500193719

Gilchrist, I., North, A., & Hood, B. (2001). Is Visual Search Really Like Foraging? *Perception*, *30*(12), 1459–1464. Retrieved from http://pec.sagepub.com/content/30/12/1459.short

Godwin, H. J., Liversedge, S. P., Kirkby, J. A., Boardman, M., Cornes, K., & Donnelly, N. (2015). The influence of experience upon information-sampling and decision-making behaviour during risk assessment in military personnel. *Visual Cognition*, *23*(4), 415–431. http://doi.org/10.1080/13506285.2015.1030488

Godwin, H. J., Menneer, T., Cave, K. R., Helman, S., Way, R. L., & Donnelly, N. (2010). The impact of Relative Prevalence on dual-target search for threat items from airport X-ray screening. *Acta Psychologica*, *134*(1), 79–84. http://doi.org/10.1016/j.actpsy.2009.12.009

Godwin, H. J., Menneer, T., Riggs, C. A., Cave, K. R., & Donnelly, N. (2015). Perceptual failures in the selection and identification of low-prevalence targets in relative prevalence visual search. *Attention, Perception, & Psychophysics*, *77*(1), 150–159. http://doi.org/10.3758/s13414-014-0762-8

Goldstone, R. L., Roberts, M. E., & Gureckis, T. M. (2008). Emergent processes in group behavior. *Current Directions in Psychological Science*, *17*(1), 10–15. http://doi.org/10.1111/j.1467-8721.2008.00539.x

Henderson, J. (2011). Eye movements and scene perception. In *The Oxford Handbook of Eye Movements* (pp. 593–606).

Henderson, J., & Ferreira, F. (2004). Scene Perception for Psycholinguists. In *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology Press.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review Psychology*, *50*(1), 243–271. Retrieved from http://www.annualreviews.org/doi/abs/10.1146/annurev.psych.50.1.243

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review*, *16*(5), 850–856. http://doi.org/10.3758/PBR.16.5.850

Henderson, J., Weeks Jr, P., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 210. Retrieved from http://psycnet.apa.org/journals/xhp/25/1/210/

Hills, T. T., Kalff, C., & Wiener, J. M. (2013). Adaptive Lévy Processes and Area-Restricted Search in Human Foraging. *PLoS ONE*, *8*(4), e60488. http://doi.org/10.1371/journal.pone.0060488

Hogg, M., & Vaughan, G. (2013). *Social Psychology* (Seventh Ed). Harlow, UK: Pearson Education Limited.

Hout, M. M. C., Walenchok, S. S. C., Goldinger, S. S. D., & Wolfe, J. J. M. (2015). Failures of Perception in the Low-Prevalence Effect : Evidence From Active and Passive Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(4), 977–994. http://doi.org/10.1037/xhp0000053

Howard, C. J., Troscianko, T., & Gilchrist, I. D. (2010). Eye-response lags during a continuous monitoring task. *Psychonomic Bulletin & Review*, *17*(5), 710–717. http://doi.org/10.3758/PBR.17.5.710

Hu, C. H., Kundel, H. L., Nodine, C. F., Krupinski, E. A., & Toto, L. C. (1994). Searching for bone fractures: A comparison with pulmonary nodule search. *Academic Radiology*, *1*(1), 25–32. http://doi.org/10.1016/S1076-6332(05)80780-9

Hübner, R. (2001). A formal version of the Guided Search (GS2) model. *Perception & Psychophysics*, *63*(6), 945–951. Retrieved from http://link.springer.com/article/10.3758/BF03194514

Ishibashi, K., Kita, S., & Wolfe, J. M. (2012). The effects of local prevalence and explicit expectations on search termination times. *Attention, Perception & Psychophysics*, *74*(1), 115–23. http://doi.org/10.3758/s13414-011-0225-4

Ishihara, S. (1917). *Test for Colorblindness*. Tokyo: Hongo Harukichi.

Itti, L., & Koch, C. (2000). A Saliency-Based Search Mechanism for Overt and Covert Shifts of Visual Attention. *Vision Research*, *40*(10), 1489–1506. Retrieved from http://www.sciencedirect.com/science/article/pii/S0042698999001637

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203. Retrieved from http://www.nature.com/nrn/journal/v2/n3/abs/nrn0301_194a.html

Jackson, T. H., Mitroff, S. R., Clark, K., Proffit, W. R., Lee, J. Y., & Nguyen, T. T. (2013). Face symmetry assessment abilities : Clinical. *American Journal of Orthodontics and Dentofacial Orthopedics*, *144*(5), 663–671. Retrieved from http://www.sciencedirect.com/science/article/pii/S0889540613007269

Jiang, Y. V, Swallow, K. M., & Capistrano, C. G. (2013). Visual search and location probability learning from variable perspectives. *Journal of Vision*, *13*(6), 13–13. http://doi.org/10.1167/13.6.13

Jiang, Y. V, Won, B.-Y., Swallow, K. M., & Mussack, D. M. (2014). Spatial reference frame of attention in a large outdoor environment. *Journal of Experimental Psychology. Human Perception and Performance*, *40*(4), 1346–57. http://doi.org/10.1037/a0036779

Kalff, C., Hills, T., & Wiener, J. M. (2010). Human foraging behavior : A virtual reality investigation on area restricted search in humans. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 168–173). Retrieved from http://www.jan-wiener.de/publications/Kalff_Hills_Wiener_CogSci-2010.pdf

Karau, S., & Williams, K. (1993). Social Loafing: A Meta-Analytic Review nd Theoretical Integration. *Journal of Personality and Social Psychological*, *65*(4), 681–706. Retrieved from http://psycnet.apa.org/journals/psp/65/4/681/

Keech, T., & Resca, L. (2010). Eye movement trajectories in active visual search: Contributions of attention, memory, and scene boundaries to pattern formation. *Attention, Perception, & Psychophysics*, *72*(1), 114–141. Retrieved from http://link.springer.com/article/10.3758/APP.72.1.114

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147. http://doi.org/10.1016/S1364-6613(00)01452-2

Koester, R., Chiacchia, K., & Twardy, C. (2014). Use of the Visual Range of Detection to Estimate Effective Sweep Width for Land Search and Rescue Based On 10 Detection Experiments in North America. *Wilderness & Environmental Medicine*, *25*(2), 132–142. Retrieved from http://www.sciencedirect.com/science/article/pii/S1080603213002664

Koopman, B. O. (1946). *Search and Screening, Operations Evaluation Group Report 56*. Center for Naval Analysis, Alexandria, Virginia.

Koopman, B. O. (1980). *Search and screening: general principles with historical applications.* New York: Pergamon Press.

Kristjánsson, Á., Jóhannesson, Ó. I., & Thornton, I. M. (2014). Common attentional constraints in visual foraging. *PLoS ONE*, *9*(6), e100752. http://doi.org/10.1371/journal.pone.0100752

Krupinski, E. (2000). The importance of perception research in medical imaging. *Radiation Medicine*, *18*(6), 329–334. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.114.1045&rep=rep1&type=pdf

Krupinski, E. a. (1996). Visual scanning patterns of radiologists searching mammograms. *Academic Radiology*, *3*(2), 137–144. http://doi.org/10.1016/S1076-6332(05)80381-2

Krupinski, E. A. (2010). Current perspectives in medical image perception. *Attention,* *Perception & Psychophysics*, *72*(5), 1205–17. http://doi.org/10.3758/APP.72.5.1205

Krupinski, E. A., Tillack, A. A., Richter, L., Henderson, J. T., Bhattacharyya, A. K., Scott, K. M., … Weinstein, R. S. (2006). Eye-movement study and human performance using telepathology virtual slides. Implications for medical education and differences with experience. *Human Pathology*, *37*(12), 1543–1556. http://doi.org/10.1016/j.humpath.2006.08.024

Kundel, H. L., & La Follette, P. S. (1972). Visual search patterns and experience with radiological images. *Radiology*, *103*(3), 523–528. http://doi.org/10.1148/103.3.523

Kundel, H. L., & Nodine, C. F. (1983). A visual concept shapes image perception. *Radiology*, *146*(2), 363–368. http://doi.org/10.1148/radiology.146.2.6849084

Kundel, H. L., & Nodine, C. F. (2004). Modeling visual search during mammogram viewing. *Medical Imaging 2004*, (pp. 110-115).International Society for Optics and. http://doi.org/10.1117/12.538063

Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic component of image perception in mammogram interpretation: gaze-tracking study. *Radiology*, *242*(2), 396–402. http://doi.org/10.1148/radiol.2422051997

Kundel, H., Nodine, C., & Carmody, D. (1978). Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Investigative Radiology*, *13*(3), 175–181. Retrieved from http://journals.lww.com/investigativeradiology/Abstract/1978/05000/Visual_Scanning,_Pattern_Recognition_and.1.aspx

Kundel, H., Nodine, C., & Krupinski, E. (1989). Searching for Lung Nodules: Visual Dwell Indicates Locations of False-Positive and False-Negative Decisions. *Investigative Radiology*, *24*(6), 472–478. Retrieved from http://journals.lww.com/investigativeradiology/Abstract/1989/06000/Searching_for_Lung_Nodules__Visual_Dwell_Indicates.12.aspx

Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature Neuroscience*, *3*(12), 1340–5. http://doi.org/10.1038/81887

Lau, J. S. H., & Huang, L. (2010). The prevalence effect is determined by past experience, not future prospects. *Vision Research*, *50*(15), 1469–1474. http://doi.org/10.1016/j.visres.2010.04.020

Lee, K., Buxton, H., & Feng, J. (2005). Cue-guided search: a computational model of selective attention. *IEEE Transactions on Neural*, *16*(4), 910–924. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1461433

Liversedge, S., & Findlay, J. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, *4*(1), 6–14. http://doi.org/http://dx.doi.org/10.1016/S1364-6613(99)01418-7

MacArthur, R. H., & Pianka, E. R. (1966). On optimal use of a patchy environment. *The American Naturalist*, *100*(916), 603–609. Retrieved from http://www.jstor.org/stable/2459298

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology Press.

Mannan, S. K., Ruddick, K. H., & Wooding, D. S. (1996). The relationship between the location of spatial features and those of fixations made during visual examination of birefly presented images. *Spatial Vision*, *10*(3), 165–188. Retrieved from http://booksandjournals.brillonline.com/content/journals/10.1163/156856896x00123

Mannan, S., Ruddock, K., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision*, *9*(3), 363–386. http://doi.org/10.1163/156856895X00052

Manning, D.J., Gale, A., Krupinski, E. A. (2005). Perception research in medical imaging. *The British Journal of Radiology*, *78*, 1–3. Retrieved from http://www.birpublications.org/doi/full/10.1259/bjr/72087985

McNamara, J. (1982). Optimal patch use in a stochastic environment. *Theoretical Population Biology*, *21*(2), 269–288. http://doi.org/10.1016/0040-5809(82)90018-1

Menneer, T., Barrett, D. J. K., Phillips, L., Donnelly, N., & Cave, K. R. (2007). Costs in Searching for Two Targets: Dividing SearchAcross Target Types Could Improve Airport Security Screening. *Applied Cognitive Psychology*, *21*(7), 915–932. http://doi.org/10.1002/acp.1305

Menneer, T., Cave, K. R., & Donnelly, N. (2009). The cost of search for multiple targets: effects of practice and target similarity. *Journal of Experimental Psychology. Applied*, *15*(2), 125–139. http://doi.org/10.1037/a0015331

Menneer, T., Donnelly, N., Godwin, H. J., & Cave, K. R. (2010). High or low target prevalence increases the dual-target cost in visual search. *Journal of Experimental Psychology: Applied*, *16*(2), 133–144. http://doi.org/10.1037/a0019569

Menneer, T., Stroud, M. J., Cave, K. R., Li, X., Godwin, H. J., Liversedge, S. P., & Donnelly, N. (2012). Search for two categories of target produces fewer fixations to target-color items. *Journal of Experimental Psychology: Applied*, *18*(4), 404–18. http://doi.org/10.1037/a0031032

Mitroff, S. R., & Biggs, A. T. (2014). The ultra-rare-item effect: visual search for exceedingly rare items is highly susceptible to error. *Psychological Science*, *25*(1), 284–9. http://doi.org/10.1177/0956797613504221

Mitroff, S. R., Biggs, A. T., & Cain, M. S. (2015). Multiple-Target Visual Search Errors: Overview and Implications for Airport Security. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 121–128. http://doi.org/10.1177/2372732215601111

Mozer, M., & Baldwin, D. (2007). Experience-guided search: A theory of attentional control. *Advances in Neural Information Processing Sysems*, 1033–1040. Retrieved from http://machinelearning.wustl.edu/mlpapers/paper_files/NIPS2007_351.pdf

Neider, M. B., & Zelinsky, G. J. (2006). Scene tonctex guides eye movements during visual search. *Vision Research*, *46*, 614–621. Retrieved from http://www.sciencedirect.com/science/article/pii/S0042698905004505

Newell, R. R., Chamberlain, W. E., & Rigler, L. (1954). Descriptive classification of pulmonary shadows. A revelation of unreliability in the roentgenographic diagnosis of tuberculosis. *American Review of Tuberculosis*, *69*, 566–584. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13148516

Nodine, C. C. F., Kundel, H. H. L., Lauver, S. S. C., Toto, L. C. L. L. C., Calvin F. Nodine, Harold k. Kundel, … Toto. (1996). Nature of Expertise in Searching Mammograms for breast masses. *Academic Radiology*, *3*(12), (pp. 89-94).International Society for Optics and P. http://doi.org/10.1016/S1076-6332(96)80032-8

Nodine, C., Mello-Thoms, C., Kundel, H. L., & Weinstein, S. P. (2002). Time Course of Perception and Decision Making During Mammographic Interpretation. *American Journal of Roentgenology*, *179*(4), 917–923. http://doi.org/10.2214/ajr.179.4.1790917

Oliva, A., Torralba, A., Castelhano, M., & Henderson, J. M. (2003). Top-down control of visual attention in object detection. In *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)* (Vol. 1, p. I-253-6). http://doi.org/10.1109/ICIP.2003.1246946

Parkhurst, D., Law, K., & Niebur, E. (2002). Modelling the role of salience in the allocation of visual selective attention. *Vision Research*, *42*(1), 107–123. Retrieved from http://www.sciencedirect.com/science/article/pii/S0042698901002504

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance: Control of Language Processes*, *32*, 531–556. http://doi.org/10.1162/jocn.1991.3.4.335

Pyke, G., Pulliam, H., & Charnov, E. (1977). Optimal foraging: a selective review of theory and tests. *Quarterly Review of Biology*, 137–154. Retrieved from http://www.jstor.org/stable/2824020

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. http://doi.org/10.1080/17470210902816461

Rayner, K., & Fisher, D. L. D. (1987). Letter processing during eye fixations in visual search. *Perception and Psychophysics*, *42*(1), 87–100. http://doi.org/10.3758/BF03211517

Rayner, K., & Pollatsek, A. (1997). the Eye-Hand Span , Eye Movements , and the Perceptual Span During of Music. *Current Directions in Psychological Science*, *6*(2), 49–53. Retrieved from http://www.jstor.org/stable/20182443

Reinagel, P., & Zador, A. M. (1999). Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems*, *10*(4), 341–350. http://doi.org/10.1088/0954-898X/10/4/304

Reingold, E. M., & Sheridan, H. (2011). Eye movements and visual expertise in chess and medicine. In *Oxford handbook on eye movements* (pp. 528–550).

Robe, R. Q., & Frost, J. R. (2002). A Method for Determining Effective Sweep Widths For Land Searches, 125. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA513088

Ruddle, R. A., & Lessels, S. (2006). For efficient navigational search, humans require full physical movement, but not a rich visual scene. *Psychological Science*, *17*(6), 460–465. http://doi.org/10.1111/j.1467-9280.2006.01728.x

Schwark, J. D., MacDonald, J., Sandry, J., & Dolgov, I. (2013). Prevalence-based decisions undermine visual search. *Visual Cognition*, *21*(5), 541–568. http://doi.org/10.1080/13506285.2013.811135

Smith, A. D., Gilchrist, I. D., Cater, K., Ikram, N., Nott, K., & Hood, B. M. (2008). Reorientation in the real world: The development of landmark use and integration in a natural environment. *Cognition*, *107*(3), 1102–1111. http://doi.org/10.1016/j.cognition.2007.10.008

Smith, A. D., Hood, B. M., & Gilchrist, I. D. (2008). Visual search and foraging compared in a large-scale search task. *Cognitive Processing*, *9*(2), 121–126. http://doi.org/10.1007/s10339-007-0200-0

Snellen, H. (1862). *Test-types for the determination of the acuteness of vision*. PW van de Weijer. Retrieved from http://dspace.library.uu.nl/handle/1874/328592

Solman, G. J. F., Cheyne, J. A., & Smilek, D. (2012). Found and missed: Failing to recognize a search target despite moving it. *Cognition*, *123*(1), 100–118. http://doi.org/10.1016/j.cognition.2011.12.006

Solman, G. J. F., & Kingstone, A. (2014). Balancing energetic and cognitive resources: Memory use during search depends on the orienting effector. *Cognition*, *132*(3), 443–454. http://doi.org/10.1016/j.cognition.2014.05.005

Stephens, D. W., & Krebs, J. (1986). *Foraging Theory*. Princeton University Press.

Strauss, B. (2002). Social facilitation in motor tasks : a review of research and theory. *Psychology of Sport and Exercise*, *3*, 237–256. Retrieved from http://www.sciencedirect.com/science/article/pii/S146902920100019X

Stroud, M. J., Menneer, T., Cave, K. R., & Donnelly, N. (2012). Using the dual-target cost to explore the nature of search target representations. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 113–122. http://doi.org/10.1037/a0025887

Stroud, M. J., Menneer, T., Cave, K. R., Donnelly, N., & Rayner, K. (2011). Search for multiple targets of different colors : Misguided eye movements reveal a reduction of color selectivity. *Applied Cognitive Psychology*, *25*(6), 971–982. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/acp.1790/full

Team, R. C. (2016). R: A language and environment for statistical computinge. Vienna,

Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/
Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology*, *40*(2), 201–237. Retrieved from http://www.tandfonline.com/doi/pdf/10.1080/02724988843000104
Treisman, A. (1991). Search , Similarity , and Integration of Features Between and Within Dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(3), 652–676. Retrieved from http://psycnet.apa.org/journals/xhp/17/3/652/
Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. http://doi.org/10.1016/0010-0285(80)90005-5
Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*(1), 15–48. http://doi.org/10.1037/0033-295X.95.1.15
Treisman, A., & Sato, S. (1990). Conjunction Search Revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(3), 459–478. Retrieved from http://psycnet.apa.org/journals/xhp/16/3/459/
Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. *Attention and Performance*, *6*, 333–361. Retrieved from https://scholar.google.co.uk/scholar?q=Treisman%252C+Sykes%252C+%2526+Gelade%252C+1977&btnG=&hl=en&as_sdt=0%252C5#1
Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *American Journal of Psychology*, *9*(4), 507–533. http://doi.org/10.2307/1412188
Tsotsos, J. (2011). *A computational perspective on visual attention*. MIT Press. Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=eT7dTizZfrYC&oi=fnd&pg=PP1&dq=tsotsos+2011&ots=LDmeCHFIF1&sig=15XIeFiJlbx6UnkVI722MAnhnoI
Tuddenham, W., & Calvert, W. (1961). Visual Search Patterns in Roentgen Diagnosis 1. *Radiology*, *76*, 255–256. Retrieved from http://pubs.rsna.org/doi/pdf/10.1148/76.2.255
Tuddenham, W. J. (1962a). Visual Search, Image Organization, and Reader Error in Roentgen Diagnosis. *Radiology*, *78*(5), 694–704. http://doi.org/10.1148/78.5.694

Tuddenham, W. J. (1962b). Visual Search, Image Organization, and Reader Error in Roentgen Diagnosis: Studies of the Psychophysiology of Roentgen Image Perception Memorial Fund Lecture 1. *Radiology*, *78*(5), 694–704. http://doi.org/10.1148/78.5.694

Vidyasagar, T. (1999). A neuronal model of attentional spotlight: parietal guiding the temporal. *Brain Research Reviews*, *30*(1), 66–76. Retrieved from http://www.sciencedirect.com/science/article/pii/S0165017399000053

Wolfe, J. (2007). Guided Search 4.0. In *Integrated Models of Cognitive Systems* (pp. 99–119).

Wolfe, J., Cave, K., & Franzel, S. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419. Retrieved from http://psycnet.apa.org/journals/xhp/15/3/419/

Wolfe, J., Evans, K., Drew, T., Aizenman, A., & Josephs, E. (2016). HOW DO RADIOLOGISTS USE THE HUMAN SEARCH ENGINE? *Radiation Protection Dosimetry*, *169*(1–4), 24–31. Retrieved from http://rpd.oxfordjournals.org/content/169/1-4/24.short

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

Wolfe, J. M. (1998). What Can 1 Million Trials Tell Us About Visual Search? *Psychological Science*, *9*(1), 33–39. http://doi.org/10.1111/1467-9280.00006

Wolfe, J. M. (2012). Saved by a Log: How Do Humans Perform Hybrid Visual and Memory Search? *Psychological Science*, *23*(7), 698–703. http://doi.org/10.1177/0956797612443968

Wolfe, J. M. (2013). When is it time to move to the next raspberry bush? Foraging rules in human visual search. *Journal of Vision*, *13*(3), 10. http://doi.org/10.1167/13.3.10

Wolfe, J. M. (2014). Approaches to visual search: Feature integration theory and guided search. In A. C. Nobre & S. Kastner (Eds.), *The Oxford Handbook of Attention* (pp. 11–55). New York: Oxford University Press.

Wolfe, J. M., Aizenman, A. M., Boettcher, S. E. P., & Cain, M. S. (2015). Hybrid foraging search: Searching for multiple instances of multiple types of target. *Vision Research*, *119*, 50–59. http://doi.org/10.1016/j.visres.2015.12.006

Wolfe, J. M., & Gancarz, G. (1996). Guided search 3.0 basic and clinical applications of vision science. *Basic and Clinical Applications of Vision Science, V. Lakshminarayanan, Ed. Dordrecht, Netherlands: Kluwer Academic*, 189–192.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Cognitive psychology: rare items often missed in visual searches. *Nature*, *435*(7041), 439–440. http://doi.org/10.1038/435439a

Wolfe, J. M., Horowitz, T. S., Van Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology: General*, *136*(4), 623–638. http://doi.org/10.1037/0096-3445.136.4.623

Wolfe, J. M., & Van Wert, M. J. (2010). Varying Target Prevalence Reveals Two Dissociable Decision Criteria in Visual Search. *Current Biology*, *20*(2), 121–124. http://doi.org/10.1016/j.cub.2009.11.066

Wolpert, D., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 593–602. http://doi.org/10.1098/rstb.2002.1238

Won, B.-Y., Lee, H. J., & Jiang, Y. V. (2015). Statistical learning modulates the direction of the first head movement in a large-scale search task. *Attention, Perception, & Psychophysics*, *77*(7), 2229–2239. http://doi.org/10.3758/s13414-015-0957-7

Yarbus, A. L. (1967). *Eye movements and vision (B. Haigh, Trans.)*. New York: Plenum Press.

Zelinsky, G. J., & Sheinberg, D. L. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(1), 244–262. http://doi.org/10.1037/0096-1523.23.1.244

Zhang, J., Gong, X., Fougnie, D., & Wolfe, J. M. (2015). Using the past to anticipate the future in human foraging behavior. *Vision Research*, *111*, 66–74. http://doi.org/10.1016/j.visres.2015.04.003