The development and trial of systematic visual search; a novel training method designed to improve the observation of workplace hazards during visual inspections conducted for risk assessment and safety auditing purposes.

A thesis submitted to the University of Dublin, Trinity College,
for the Degree of Doctor of Philosophy by;

Victor Hrymak

School of Nursing and Midwifery
Faculty of Health Sciences
University of Dublin
Trinity College
Supervisors;
Dr Jan De Vries & Dr Maria Chiara Leva
April 2018

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Signed__________________________ Date: _____________________

Victor Hrymak

Abbreviations

EHS; Environmental Health and Safety

SVS; Systematic Visual Search

VIP; Visual Inspection Performance

OSH; Occupational Safety & Health

Acknowledgements

I would like pay tribute and state my sincere thanks to the following individuals who have been responsible for allowing this PhD to be completed.

To Dr Jan De Vries, for agreeing to supervise in the first instance, and thereby allowing this research to proceed to completion. Without Dr De Vries's patience, direction and enthusiasm, this PhD would have remained a distant aspiration.

To Dr Chiara Leva for invaluable support and advice, as a colleague and second supervisor.

To the School of Nursing and Midwifery, Faculty of Health Science, Trinity College Dublin, who provided the infrastructure and administrative support for the research.

To Dr James Curtin, the Head of School of Food Science & Environmental Health and Professor Michael Devereux the Director of the College of Science and Health in the Dublin Institute of Technology. Their very generous funding of all fees, as well as unconditional logistical support, facilitated the production of this thesis.

To Dr Detta Melia, Gereva Hackett, Julianne Forrestal and Rachel Beirne for their crucial role, in securing participants. To Dr's; Natasha McCormack and Ciara Walsh, for their invaluable role, in the running of the trials.

And finally, but most importantly; to my wife and five young children.

Summary

Accurate visual inspection performance, is paramount for the legally required risk assessment of workplaces conducted by Environmental Health and Safety professionals. However, visual inspections are problematical when observable hazards that should be seen, are missed. Any resulting unobserved hazards, can thereby lead to accidents, ill health and financial loss.

A primary reason for these unobserved hazards, is the lack of standardisation allowing idiosyncratic visual inspection practice which in turn, can lead to incomplete visual searches. Risk assessment methodologies have seen many advances since their widespread introduction into the European regulatory landscape in the late 1980's. In sharp contrast, visual inspection accuracy has been largely neglected by the academic and professional Environmental Health and Safety community.

To address this problem, a novel method called systematic visual search was developed and tested under ecologically valid, randomised controlled trial conditions. Two studies were conducted; a kitchen safety and hygiene inspection task, and a light aircraft pre-flight inspection task. Both studies involved visual inspections being evaluated by ascertaining the mean percentage of observable hazards seen during a thirty minute inspection period.

The aim of systematic visual search is to ensure an exhaustive observation of all areas under analysis. This requires that the inspection space is sub-divided into constituent elements. Each element is then individually selected and subjected to an exhaustive visual search, using a pre-scribed left to right eye scanning pattern, similar to reading a book.

In the first study, participants (N = 211), previously trained in occupational and food safety, visually inspected industry standard kitchens. The experimental group (N=107) received training in systematic visual search. The control group (N=104) conducted inspections as per their normal custom and practice. The kitchens contained a known number of observable hazards, representative of workplace conditions that such professionals normally encounter.

Twelve randomised controlled trials were conducted using five industry standard kitchens. These kitchens had a precise number of observable hazards, a proportion of which were "planted" to achieve "real world" conditions. Participants were assigned to a control or experimental condition. In the control condition, they were instructed to use their customary visual inspection practice. In the experimental condition, participants received training in

systematic visual search before performing a visual inspection on the same kitchen. All participants wrote down hazards observed, which were then collected for data analysis.

The aim of the aircraft pre-flight inspection task was to see if any beneficial effect in visual inspection performance demonstrated in the kitchen study, could be reproduced in a highly technical industrial inspection scenario. Systematic visual search was therefore tested amongst an experienced cohort of aviation maintenance engineers (N=26). In addition, the systematic visual search educational pedagogy in this study was delivered by a senior aviation engineer. This allowed any bias from researcher's involvement in the first study, to be eliminated as well as confirming further potential industrial applicability. Twenty six experienced aviation maintenance engineers conducted a visual inspection on an actual light aircraft. Again the aircraft had a precisely known number of hazards and "plants". After randomisation, thirteen aviation maintenance engineers were asked to conduct a pre-flight check using their customary visual inspection practice. The remaining thirteen participants were taught systematic visual search by the senior engineer before visually inspecting the aircraft using this method. The results from both studies are shown below.

Kitchen Study	**N**	**Mean % hazards observed**	**SD**	**95% CI's**	***p* value**	**Cohen's *d***
Control	107	32.96	9.02	[31.24- 34.70]		
Experimental	104	49.64	10.88	[47.53- 51.76]	≤.001	1.84
Aviation Study						
Control	13	37.90	9.65	[32.09-43.76]		
Experimental	13	63.66	10.11	[57.54-69.77]	≤.001	2.66

The results demonstrated the higher visual inspection performance of systematic visual search, compared to current inspection custom and practice. Even so, visual inspection performance was below expectation, as many observable hazards were simply not seen. The literature details the many reasons for these generally low results, concluding that visual inspection is an error prone task and difficult to do well. Therefore the practice of visual inspection needs further academic attention, in order to improve workplace safety. In this respect, systematic visual search has a potentially important role to play.

TABLE OF CONTENTS

List of Figures

List of Tables

CHAPTER 1 INTRODUCTION

1.01 Chapter Overview

This introduction chapter begins by stating the problem addressed by this thesis; not seeing observable hazards during workplace inspections. This is followed by a summary of current custom and practice for visual inspection which is the fundamental process used to observe workplace hazards during inspections. The chapter will continue by looking at the role and importance of visual inspection, the guidance for its conduct and the consequences of visual inspection failures. Furthermore, an outline will be presented of efforts to improve visual inspection, including the intervention used in this thesis, systematic visual search (SVS). This chapter will then conclude by presenting the thesis aims, objectives, hypothesis and research question. It should be noted here that the intervention itself consisted of three separate components; the training in the systematic visual search method, the lecture slides used in the pedagogy and a nine page handout given to experimental group participants to facilitate them in writing down observed hazards (see section 4.2.4)

1.1 Setting Out a Context for this Thesis

Each day will bring thousands of workplace inspections conducted by EHS professionals who want to decide on just how safe a particular workplace is. This is a legal requirement that is common to all advanced market economies. The reason for requiring workplace inspections, is to see what measures are required to prevent accidents, ill health and financial loss (HSA, 2015). These inspections and the legal risk assessment requirements it underpins, has long been part of the European Union's social protection role stretching back to 1950's. The conduct of these visual inspections varies greatly. This is reflected in a bewildering array of descriptors including; hazard identification, risk assessment, risk audits, risk reviews, risk surveys, safety audits, safety inspections, safety reviews, safety checks, safety tours, safety visits, safety surveys, safety walkthroughs, health and safety visits, health and safety audits, due diligence visits, accident investigations, inspections, surveys, scoping visits, familiarisation exercises, management safety walkabouts and forensic investigations. Even though these descriptors are commonly used, they are interpreted differently by EHS professionals. However they all possess one commonality, that of observation. In each and every workplace inspection, regardless of its title, aim or provenance, someone will have used their eyes to look around the workplace for defects, hazards or targets. No one is advised or expected to conduct such inspections blindfolded!

Even though visual inspection is a fundamental requirement for the risk assessment process, there has been very little research as to the effectiveness of inspections conducted for workplace safety by EHS professionals (Woodcock, 2014). Furthermore when empirically tested in the field, the limitations of visual inspection are demonstrated. For example, the visual inspection performance (VIP) of 37 U.S. bridge inspectors was found to be incorrect in 56% of their inspections conducted (Moore et al, 2001).

1.2 Prevalence & Reasons for Visual Inspection

Risk assessments are a legal requirement for virtually all organisations in the EU (HSA, 2015). The sheer number of risk assessments and by extension, visual inspections required in a modern economy are therefore vast. Scholars report that a large number of “real world” visual inspections are conducted on a daily basis (Biggs & Mitroff, 2013). The scale of this requirement can be gauged by considering that in 2011, the number of Irish active enterprises (defined as businesses registered with the Irish Companies Registration Office) was over 185,000 (Business in Ireland, 2012). As a minimum therefore, there was a requirement for 185,000 risk assessments and thereby the same number of visual inspections, to be carried out at least once in Ireland. A multiple of this number will be the true figure of risk assessments required, given the legal obligation to review and re-write assessments as circumstances change (HSA, 2015).

The number of visual inspections conducted becomes far larger when considering risk assessment on a national and international scale. A recent study by the European Agency for Occupational Safety and Health (EU-OSHA, 2014) shows just how widespread and extensive risk assessment practice is. Regarding the EU-27 it was found that 72% of enterprises, reported risk assessments or workplace checks were being carried out at regular intervals. Therefore, taking the EU-27 as a whole, there are circa 218 million European workers that depend on risk assessment for their workplace safety (EU-OSHA, 2014). A further impetus for workplace risk assessment, and again for visual inspection, has been the growth in the quality movement and the use of international standards such as BS EN 14000: (2004); Environmental Management, ISO 45001: (2018) Occupational Health and Safety Management Systems, and ISO 27000: (2005) Food Safety Management. These standards are now established and consensus documents designed to manage specific risks (Wadsworth & Walters, 2014).

1.2.1 The Human and Economic Importance of Visual Inspection

In addition to the legal obligation to ensure workplace safety, there are further important socio-economic reasons to explain why EHS professionals conduct so many visual inspections. They are conducted for example to ensure public safety. Visual inspectors are responsible for tasks of national and social importance that include x-ray security screening, aircraft inspection, industrial inspection and radiology. It has been stated that these “real world” visual inspections have life or death implications, should any relevant defects hazards or targets not be observed, (Biggs & Mitroff, 2013b; Mitroff, et al 2015; Nakayama, & Martini, 2011; Wolfe et al, 2005).

For example in aviation, visual inspection is the most frequently used technique accounting for some 80% of all inspections used in the sector (Drury & Watson, 2002). These scholars go on to state that the continuing airworthiness of the civil aviation fleet depends on visual inspection reliability. They also state that 18% of aircraft incidents and accidents can be attributed to visual inspection failures. (Drury, 2005; Melloy et al, 2000; Asada et al, 1998).

1.2.2 The Legal Requirement for the Visual Inspection of Workplaces

The main driver for EHS professionals conducting visual inspections, comes from the legal obligation for risk assessment. This requirement extends to the vast majority of EU businesses and organisations (Wadsworth & Walters, 2014). There is only one small jurisdictional exemption to this duty. In the UK, those organisations with less than five employees are exempt from written risk assessments (HSE, 2017). Therefore, the legal and professional practice of risk assessment, which has at its very centre, a visual inspection component, is an accepted and very widely used concept in advanced market economies (Wadsworth & Walters, 2014). To explain further, risk assessment is a two-step process. It firstly requires a hazard identification to be conducted, which is where visual inspection becomes involved. Once an evaluation of these identified hazards has occurred, then the risk assessment is complete (HSA, 2015). The risk assessment concept together with its fundamental visual inspection component has been long extended to the regulation of food safety, fire safety and environmental protection (Wadsworth & Walters, 2014).

It should be noted here, that there are further specialised risk assessment methodologies that are used extensively by process industries, particularly at the design stages of plant and equipment. These methodologies are not the subject of this thesis. Such methodologies include HAZOP, FMEA, ETA & FTA. A full listing and a brief explanation of these methods is given by Mariken et al (2013).

1.2.3 Overarching Causes of Visual Inspection Failures.

The fundamental problem that causes so much difficulty for EHS professionals is defect, hazard or target location uncertainty (Cain et al, 2014; Eckstein, 2011; Mitroff et al, 2015; Wolfe et al, 2016). Simply put, EHS professionals cannot possibly know in advance the frequency and location of all the workplace hazards they are sent to look for. So unless there is a strictly proceduralised method as to how to look, where to look, and for how long to look, such professionals tend to rely on their own idiosyncratic custom and practice. Unfortunately, a result of this subjectivity is that the duration and nature of their visual search behaviour is rarely connected to the number of hazards within the area under analysis.

On top of this fundamentally limiting factor, there are further causes of visual search error that detrimentally affect visual inspection performance. It should be noted here that not seeing observable defects, hazards or targets is referred to as error by visual search scholars (Eckstein, 2011) and this phrasing will be used throughout this thesis. The causes of these many and varied visual search errors can be categorised under the headings of sensory-perceptual (for example, in-attentional blindness), cognitive bias (not regarding the object seen as a hazard) and organisational (not allowing enough time for the inspection), Cain et al, 2014; Eckstein, 2011; Mitroff et al, 2015; Wolfe et al, 2016). A full description of these causes of visual search error will be given in Chapter 2 (Theoretical Underpinnings) and are summary listed in Table 3.

The evidence for EHS professionals, not observing observable defects, hazards or targets has been well reported. For example See, (2013) summarises that visual inspection performance (VIP) for the US industrial engineering sector, resides somewhere between 9% and 100% accuracy. This is a very large range for a widely used observational task that intrinsically requires accuracy. As Chapter 3 (Visual Inspection Literature) will show, this range effectively means that for some observational tasks, no objects of interest are seen, and in others, all of them are seen. Even in highly regulated industries such as the US nuclear industry See, (2015) reports the accuracy rate for the visual inspection of nuclear weapons components to be 85% with a mean 80% accuracy rate for this sector as a whole.

There is also the issue of attempting to describe what is current visual inspection custom practice by EHS professionals. There is some indication of current custom and practice from a survey of forty experienced Irish EHS professionals (Hrymak et al, 2015). In this small study visual inspection practice was idiosyncratic with little consensus as to what to look for, how to look, in what order to look, or how long to look for. What is thought to constitute current custom and practice, is further discussed in section 2.2.6.

1.2.4 Illustrative examples of serious visual inspection failures

The most tragic consequence of not seeing observable workplace hazards occurs when there is resultant loss of life. This is all the more regrettable, if it is shown that such fatalities could have been prevented by better performance during the pre-accident visual inspection. An extensive and judicially investigated example occurred in 2011, when a fire in a nursing home caused the deaths of fourteen residents at the Rosepark nursing home in Scotland. The coroner's report found that the cause of this fire was inappropriate storage of flammable aerosols in an electrical cabinet (Lockhart, 2011). The coroner found that this workplace hazard was not observed during a recent inspection by a health and safety consultant. The coroner concluded that; an adequate risk assessment by this EHS professional should have identified this hazard, which if rectified, would have prevented the fatalities.

The Rosepark fire, evidences how an observable workplace hazard was simply not seen by an EHS professional during his visual inspection. This case tragically illustrates the problem this thesis seeks to address. Why was a perfectly observable hazard not seen by the very professional who inspected this particular workplace with the specific intent of identifying such workplace hazards?.

There are further examples of visual inspection failures from the construction engineering, airport security screening, and medical imaging diagnostics sectors. They are included here to illustrate the wider range of scenarios and sectors, from which visual inspection failures have been demonstrated. In a study of the capability of construction workers to observe hazards, Albert et al, (2017) reported a baseline figure of only 32%. Regarding x-ray security screening for carry on hand luggage at airports, Wolfe et al, (2005) found that if 1% of the luggage passing through x ray machines contained prohibited items such as; knives, guns or bombs, the error rate of observers was 30%. Regarding medical imaging, there is a 0.3% prevalence rate for cancerous lesions presenting on scans. Mammographers looking at these scans, exhibited a 50% error rate in their observations (Evans, et al, 2013).

An interesting extension of observational hazard error from members of the general public was demonstrated in a study by Douglas et al, 2015. Here, 153 respondents who visited a science gallery with an exhibition on risk, were asked if they had noticed three simulated but fake hazards within the exhibition space. These hazards were; a damaged electrical switch box, a leaking chemical cupboard, and a leaking pipe. Recall of the chemical cupboard was the highest at 51% of the respondents with 37% and 0% recalling the switch box and leaking pipe. Only three people reported any of these hazards to exhibition staff.

In summary, visual inspections resulting in observable defects, hazards and targets not being seen, is a phenomenon that regularly occurs whenever EHS professionals go about their business. There is therefore a clear need to improve current visual inspection conduct.

1.3 Efforts to Improve Visual Inspection Conduct

The main reason why current custom and practice of visual inspection remains idiosyncratic is due to a lack of standardisation (Aven, 2011; Hrymak et al, 2015; Johanson & Rausand, 2015). It is therefore useful at this point, to summarise professional practice controls that exist for the conduct of visual inspection. These include published guidance on visual inspections, the checklist approach, the Chartered Surveyor method for building surveys, training methods and the overarching legal doctrine of professional negligence. Other than the Chartered Surveyor method, it will be seen that the guidance on visual inspection conduct remains vague, ambiguous and with little professional consensus as to conduct.

1.3.1 Guidance for Visual Inspection Conduct

The UK Labour Inspectorate Regulator has stated; "there are no fixed rules about how a risk assessment should be carried out" (HSE, 2017). This ambiguity is echoed by the vague nature of regulatory guidance from the Irish, British and US Labour Inspectorates when they advise on the conduct of risk assessments. The Irish Health and Safety Authority state that; risk assessors should "walk around the workplace and look afresh at what could reasonably be expected to cause serious harm" (HSA, 2006). The UK's Health and Safety Executive state that "to accurately identify the potential hazards in your workplace, a good starting point is to walk around your workplace and think about any hazards" (HSE, 2014). The US based

Occupational Safety and Health Administration, similarly state that their inspectors will "walk through the workplace inspecting for hazards that could lead to employee injury or illness" (OSHA, 2016). The only commonality presented in the guidance from these three Labour Inspectorates, is the suggested requirement for visual inspection as part of the identification of workplace hazards. Despite visual inspection being required by these Labour Inspectorates, guidance remains vague and as a result, professional conduct remains idiosyncratic. Furthermore, the paucity of research on visual inspection within the EHS community means that there is little evidence as to what is the best method to observe workplace hazards during workplace inspections. This point will be detailed Chapter 3 (Visual Inspection Interventions)

1.3.2 Checklist Based Inspection

The most popular risk assessment method for workplace inspections is the checklist (Clift et al, 2011). This method, which axiomatically uses the visual sense has resulted in literally hundreds of differing checklists and templates being available for identifying hazards during workplace inspections. They cover all manner of workplace hazards such as manual handling, display screens, chemicals, slips trips & falls, hygiene and general housekeeping issues (Clift et al, 2011). Some scholars are very supportive of checklists with Neathey et al, (2006) concluding that all-encompassing checklists for common risks in small and medium sized enterprises are very useful. Some "popular media" commentators also view checklists very highly. For example Gawande, (2009) writing in his New York Times bestseller; "The Checklist Manifesto" credits them with achievements varying from reducing hospital infection rates to the well known incident whereby a stricken passenger aircraft was landed on the Hudson River in New York. This author points to checklist utility stemming from their ability to inform on what tasks need to be done, including what objects need to be looked at.

But for such a ubiquitous practice, it is noteworthy that very little literature exists from the risk analysis discipline presenting any empirical evidence in support of checklists (Clift et al, 2011; Neathey et al, 2006). Furthermore, as checklists are in effect directed visual inspections, they are subject to the same sensory-perceptual, cognitive bias and organisational causes of visual search error, as detailed in Table 3 and briefly mentioned above. In addition, and as the two studies in this thesis point out, a major disadvantage of checklists is that they do not inform on how best to look for objects of interest.

1.3.3 The Chartered Surveyor Method of Visual Inspection

There is one group of professionals, who have proceduralised visual inspection conduct. Members of the UK based Royal Institute of Chartered Surveyors have professional practice guidelines, on how to conduct visual inspections (or surveys in their jargon) for building condition reports (Hollis, 2000; Reddin, 2016; RICS, 2002; RICS, 2010). These professionals need to observe all defects, hazards and targets that can affect building condition, for example damp, disrepair, subsidence etc. This group of professionals have evolved their own visual search tradition to minimise the non-observation of building related hazards. When visually inspecting a room in a building, these professionals begin by selecting a specific element of the room for exclusive observation. The selection order recommended by Chartered Surveyors profession is commonly; the ceiling first, then each wall, windows, doors and chimney breasts. Once selected, the element is the sole focus of the visual inspection and on completion of observation, is not returned to again (Hollis, 2000).

1.3.4 Training to Improve Visual Inspection Performance

Many commentators have stated the importance and positive effect on VIP from training. In an overarching comment regarding "real world" professional search practice, the number of defects, hazards and targets is unknown and the situation requiring the search cannot be changed (Eckstein, 2011). Therefore the search accuracy of the observer is arguably the weakest link and thereby, offers the greatest potential for improvement (Eckstein, 2011; Biggs & Mitroff, 2013b). Various visual search scholars state that visual inspection can be improved with training (see for example Gramopadhye, et al, 1997; Gramopadhye et al, 1997b; Wang et al, 1997; Swaninger, 2005). Large improvements in VIP to over 70%, have also been reported when using a training intervention involving a mnemonic to remember construction site hazards (Albert et al, 2017). A detailed presentation on the effect of training interventions on VIP is given in Chapter 3 (Visual Inspection Literature).

1.3.5 Using Specific Eye Scanning Patterns for Visual Inspection

At this point, it is time to introduce a further strand of visual search inquiry that has a direct and important bearing on the thesis aim, namely how should the searcher actually look. In short, what visual search behaviours should be recommended in the interest of observational accuracy. A full and detailed explanation of visual search behaviour is given in sections; 2.2.2, 2.2.3 and 2.2.4. In short, visual search scholars state that improved visual inspection performance should be exhibited by inspectors using non-random as opposed to random visual search strategies (Nalangula et al 2006: Nickles et al 2003; Wang et al, 1997). There are a small number of studies that support this idea by using a set eye scanning pattern to look for observable defects hazards or targets. Some scholars state rather axiomatically, that the only way an exhaustive visual search can be conducted, is by using a visual search behaviour which ensures that all areas are included in the inspection process (Melloy et al, 2000). Other visual search scholars state that consistency during the visual search strategy, or any process that can reduce uncertainty faced by professional searchers looking for objects of interests; will be associated with better VIP. (See, 2012; Eckstein, 2011).

To visually search in a non-random manner, the observer is expected to use a set eye scanning pattern during the visual inspection (Nalangula et al 2006: Nickles et al 2003; Wang et al, 1997). Examples of a non-random eye scanning patterns are exemplified if an object under analysis is observed using a left to right, reverse “snakes and ladders or up & down left to right or eye scanning pattern (see Figures 1, 6 & 7). Simply put and in summary; using non-random eye scanning patterns during visual inspections has been demonstrated to improve VIP (Nalangula et al 2006: Nickles et al 2003; Wang et al, 1997). A detailed account of the effect of using set eye scanning patters during visual inspections, is presented in Chapter 3 (Visual Inspection Literature).

1.4 SVS; A New Approach for Visual Inspections

Current guidance for visual inspection remains imprecise. There is an over-reliance on checklists and prior knowledge of what to look for. Given the tragic visual inspection failures listed above in section 1.2.4 (and with further examples listed in Chapter 3), it would appear to be necessary to consider additions to current top down approaches. So it is now proposed that a new approach to the conduct of visual inspection is both possible and warranted. This new approach is intended to improve two observational processes by precisely proceduralising visual search behaviour into element selection and eye scanning pattern. Firstly the order of elements selected during visual inspection, will be exactly as recommended by the Royal Institute of Chartered Surveyors, (see section 1.3.3) Secondly to improve observational accuracy itself, a set eye scanning pattern is to be specifically used, as recommended by Nalangula et al (2006); Nickles et al (2003); Wang et al, (1997).

By combining the Chartered Surveyor method for room element selection (see Hollis, 2000) and recommended eye scanning patterns from the literature (Nalangula et al 2006: Nickles et al 2003; Wang et al, 1997), a novel visual inspection method, here called systematic visual search (SVS) was conceptualised. SVS is therefore far removed and far less reliant on top down information whereby checklists or prior knowledge is the primary driver of what to look for during the visual inspection. Instead, the SVS method proceduralises visual inspection by firstly selecting only one element at a time in a room to observe, for example a wall. Then the entirety of the wall is exhaustively searched by using a set eye scanning pattern to ensure vision is directed over all areas of the wall.

1.5 Aim, Objectives, Research Hypothesis & Question

SVS represents a potential improvement to current visual inspection custom and practice by EHS professionals. This is because introducing a visual search behaviour that consists of; selecting a specific element or area, then using a set eye scanning pattern to ensure all areas of the selected inspection space are exhaustively observed, may well lead to beneficial effects on VIP in practice. SVS has until now remained empirically untested in the field. This study has been designed to make up for this and the following aims, objectives, research hypothesis, and question were therefore formulated.

The aim of this thesis is the development and trial of SVS; a method to potentially improve the observation of workplace hazards during visual inspections. In short, if SVS does have merit in terms of improving visual inspection performance, then it needed empirical evidence of any beneficial effects from testing under precisely defined field conditions. This aim led to the formulation of three objectives that together would allow for the development and trial of SVS. These objectives are now listed and explained below

Objective 1; the development of the SVS method. Here the objective was to exactly formulate what the SVS method would be. This required an exhaustive literature search of visual inspection methods before the preferred observational method could be selected for further evaluation See Chapter 3 (Visual Inspection Literature).

Objective 2; the design of a training session for the SVS method.

Once the SVS method had been defined, a training session needed to designed so that prospective users could utilise the method during their visual inspections. This objective led to a syllabus being designed together with an accompanying PowerPoint presentation (see

Appendices 1-4. This syllabus was delivered during a half hour training session that included a practical demonstration of SVS by the lead investigator (see Sections 4.2.1, 4.2.3 and 4.2.4)

Objective 3; the testing of SVS under post positivist experimental conditions.

The final objective was to see what effect SVS would have on visual inspection performance in the field. An experiment was therefore devised using real life workplaces. Recruited participants were randomly allocated in control and experimental conditions. In the control condition, participants were asked to visually inspect in their normal manner. In the experimental condition, participants were trained to use SVS (see Chapter 4 Methodology)

During the experimental design, a research hypothesis was required. However as the effect of SVS was at the time obviously unknown, its use could have increased as well as decreased visual inspection performance. Therefore the research hypothesis was phrased in a two tailed manner as follows;

"when looking for observable workplace hazards, the use of SVS will result in different visual inspection performance rates when compared to current practice".

In addition, a broader qualitative research question was formulated for this thesis as follows;

why are visual inspections often flawed and how can these flaws be overcome?

1.6 Overview of Chapter Content

This section will briefly outline the content for the remaining thesis chapters.

Chapter 2 Theoretical Underpinnings. This chapter will summarise visual search science, which is a vast area of multidisciplinary knowledge, drawn from the natural and social sciences. The reason for its inclusion is the need to comprehend the complexity of how humans see and consciously interpret objects of interest, such as defects, hazards or targets.

Chapter 3 Visual Inspection Literature. This chapter will evidence current VIP by EHS professionals. It presents those studies whereby the field based observation of defects, hazards or targets has been empirically investigated. It will also include published studies from the effects of eye scanning patterns on VIP.

Chapter 4 methodology. This chapter begins with a detailed consideration of the epistemological philosophy underpinning the methodology used and the design of the randomised controlled trials. It will also include the trial procedures used for the intervention together with a statistical analysis plan for the data generated.

Chapter 5 Results. This chapter begins by presenting comparative demographic data for control and experimental groups. A detailed analysis of the results then follows to highlight the effect of SVS on VIP using descriptive and inferential statistical tests. The chapter concludes with a comparative assessment of the kitchen and aviation studies together with qualitative comments from participants.

Chapter 6 Discussion. This chapter begins by considering the strengths and limitations of the two studies conducted. Evidence as to why SVS demonstrated better VIP compared to current custom and practice is then discussed together with speculation as to why overall VIP was found to be below expectations. The chapter concludes with a discussion of the thesis implications for the wider EHS community

Chapter 7 Conclusions and Recommendations. This chapter succinctly encapsulates the main findings from the thesis based on the study data and the wider literature. It will also present recommendations for the wider EHS community as well as detailing further research in order to improve the conduct of visual inspection. The chapter ends with a section on possible applications of SVS.

CHAPTER 2 THEORETICAL UNDERPINNINGS

2. 01 Introduction

This chapter consists of the theoretical underpinnings for the two studies. It does this from a variety of domains with relevance for the conduct of visual inspections and why these may be flawed. It will include how humans see, the role of visual inspection, together with a taxonomy on the causes of visual search error. The body of relevant literature for this thesis is substantial as already outlined in the Chapter 1 (Introduction). It comes from several theoretical domains such as social psychology, vision science, visual search, cognitive psychology, neuroscience, industrial engineering and organisational theory. However it is argued that visual search, due to its inherent complexity and wide range of applications necessarily has, and benefits from, this large body of literature. The core of relevant principles from these domains will be presented in this chapter, whilst specific empirical research by Environmental Health and Safety (EHS) professionals in the “real world” will be covered in Chapter 3 (Visual Inspection Literature).

2.1 How Humans See

Vision is our dominant sense. Vision allows us to derive most of our information about the world in terms of location motion and object recognition (Wade & Swanston 2013). These scholars further explain that vision begins with light in the form of photons from distal objects, entering the eye. These photons are firstly inverted by the lens before reaching the retina. The rods and cones in the retina convert these upside down images into electrical energy, which result in chemical communication to the visual cortex of the brain via synapses and neurons. The brain interprets these electrical impulses and thereby allows a conscious perception of the image with an understanding of what is being seen by the human (Wade & Swanston 2013).

How these sensory, perceptual and cognitive abilities result in a conscious visual experience is truly remarkable, both in terms of what is achieved, as well as the speed at which this takes place (Wade & Swanston, 2013). Consider the example of an EHS professional inspecting a workplace and looking for hazards. The complexity and sheer scale of this visual environment, in terms of the amount of photons entering the eye is unimaginably vast. The visual stimuli will include for example, the spectrum of colours available, the number of shapes, location distances and any movement from velocity, depth, size and perspective.

Even so it is remarkable to consider that it will take humans from the moment the eyes see an object, only about 25 milliseconds to observe and consciously understand what that object is (Wade & Swanston, 2013).

Scholars also report that this astonishingly quick and highly effective cognitive ability at visual object recognition continuously responds to new visual stimuli. If more visual stimuli are received for example, a fly in the workplace is observed by peripheral vision then saccadic eye movements (rapid eye flicks) and if necessary head movements will occur to direct more of the relevant photons from the fly to the fovea. The fovea is the central area of the retina that allows the greatest resolution. The automatic and rapid saccadic eye re-adjustments results in the object of interest being fixated (stared at), with improved object recognition again measured in the millisecond range. Saccades, which take about 20-30 milliseconds to complete represent the fastest physical movement humans can perform (Wade & Swanston, 2013). Saccades move at this astonishing speed because foveal vision only accounts for about 1-2 degrees of total vision. Hence rapid eye movements are required to constantly ensure maximum optical resolution for any objects of interest fixated.

The visual system processes involved in object recognition, which includes the observation of hazards, therefore allows for complex decision analysis and behaviour such as safely crossing a busy road to be carried out very effectively and on a continual basis (Nakayama, & Martini 2011). These scholars report that complex visual stimuli from for example, emotional scenes can become conscious perception in less than 100 milli-seconds, again showing very fast cognitive processing and perceptual understanding.

2.1.1 Attentional Theory

Given the sheer scale of visual stimuli continually entering the eye, the brain has evolved a mechanism to filter out unnecessary data and only allow useful visual stimuli to reach the brain for object recognition and conscious perception (Wolfe et al, 2005). This filtering is necessary, as it is simply not possible for the brain to process all the stimuli continually entering the eye. The question now arises, as to what makes humans notice for further processing in the brain, specific objects of interest from the billions and billions of visual stimuli constantly entering the eye. This multidisciplinary field is known as attention and is defined as the “sustained concentration on a specific stimulus, sensation, idea, thought or activity enabling one to use information processing systems, with limited capacity, to handle vast amounts of information available from the sense organs and memory stores" Colman, (2009).

The visual system needs the process of visual attention deployment to allow humans to simultaneously recognise visual objects of interest and their meaning in any visual environment (Wade & Swanston, 2013; Wolfe & Horowitz, 2004). These scholars both point out this attentional deployment processes begins with the filtering of this visual stimuli excess in the eye itself. Through head and eye movements, important stimuli are directed to be processed by the fovea where resolution is the clearest. Remaining visual stimuli receive less analysis to cope with the total amount that has entered the eye.

The highest level of resolution for visual stimuli occurs in the central region of the human field of view (visual lobe), which corresponds to the fovea with no more than 2 degrees to the periphery (Wade & Swanston, (2013). Moving the visual stimuli of interest away from the fovea and into the retinal periphery, results in large losses of acuity. Eyes will fixate (stare) and remain relatively still for about 20-30 milliseconds. Eyes will then flick very rapidly to the next object of interest. Outside this central region peripheral vision is still operating so that humans are aware of the images, but not with the same optical acuity.

"Some visual search tasks are easy some are not". This statement, by Wolfe & Horowitz, (2004) is due to the effect of object attributes on visual cognition. Some object attributes are better at guiding the deployment of visual attention and humans are therefore more likely to notice them. The evidence is that size, motion and colour are thought to be the best attributes to guide the deployment of visual attention. There are three ways this attentional deployment can occur, bottom up stimuli processing, top down stimuli processing and scene guidance (Wolfe et al, 2015). Bottom up stimuli processing refers to some characteristic allowing the object to stand out from its surroundings such as movement size or colour. Top down stimuli processing refers to cognitive functions, whereby the observer is trained to look for specific targets such as hazards in a workplace. Scene guidance refers to human expectation of visual scene topology. Asked to look for fish for example, observers will look to the nearest water source and not at the sky. Attentional deployment therefore allows important stimuli or objects of interest to be selectively channelled to the brain, which in turn allows object recognition and conscious perception (Wolfe & Horowitz, 2004). Therefore when considering the observation of workplace hazards, it is theorised that objects of interest are seen when visual attention is guided by top down stimuli processing, bottom up stimuli processing, or scene guidance.

A summary of the attributes or characteristics of objects, into five categories that can guide the deployment of human attention and allow an object to be recognised during a visual search are shown below in Table 1 below (from Wolfe & Horowitz, 2017). What this table shows, is those salience characteristics or observable features of an object that are thought to deploy visual attention in decreasing order of effect. The evidence for salience is therefore presented as decreasing attribute categories ranging from undoubted to probably not likely to deploy visual attention. For information, the Vernier effect refers to the movement of gradations against each other when using a Vernier caliper measurement device. Geons refer to basic shapes within larger shapes, such a lego brick in a lego wall.

Table 1 Attributes that Might Guide the Deployment of Visual Attention

Undoubted Attributes	**Probable Attributes**	**Possible Attributes**	**Doubtful Attributes**	**Probable non Attributes**
Colour	Luminance	Lighting	Novelty	Intersections
Motion	Vernier effect	Glossiness	Apha-numeric	Geons
Orientation	Depth	Number		Faces
Size	Line termination	Aspect ratio		Semantic sets eg
	Topology			Scary or animal
	Curvature			

In the context of this thesis, the table above is saying that; when EHS professionals are conducting their visual inspections, any hazards with the visual characteristics listed in column one (Undoubted Attributes) such as colour should in theory, be seen more successfully than hazards with visual characteristics form the next column (Probable Attributes) such as luminance, and the next column (Possible Attributes) etc. Applying this table to the thesis, a hazard that is large should in theory, be seen more often than a hazard that is brightly lit. A full discussion on the hazard attributes that were found in this thesis to deploy participant visual attention is presented in sections 5.1.5; 5.4.5; 6.4.1and 6.4.2.

2.1.2 Definitions; Visual Search, Visual Inspection & Visual Behaviour

The overarching academic discipline involved in defect, hazard or target observation and recognition is visual search. Visual inspection is a closely resembling sub-set of visual search and the visual inspection term, is extensively used in the industrial quality sector. Furthermore “real world” professionals use their own specific terms to describe their visual searches such as scans or examinations by medical imaging diagnosticians, checks by x-ray baggage security staff and surveys by chartered surveyors. As the academic discipline of visual psycho-physics includes social as well as natural scientists, a degree of interchangeability and similar meaning is exhibited especially when considering the terms visual search and visual inspection.

2.1.3 Visual Search Definitions

Visual search has been defined as “an experimental task, the objective of which, is to find a particular element in a visual display” (Colman 2009). It is further variously described by scholars here referenced respectively as “what it sounds like; looking for something”; “locating target objects among distractors” and “the process of locating a target or defect within an area of interest” (Wolfe, 1996; Biggs & Mitroff, 2013; Nalangula et al, 2006).

Visual search is further exemplified as something we all do in our daily lives, that the items searched for can be more or less anything, and that our environment continuously triggers visual search activity (Nakayama, & Martini 2011; Wolfe & Horowitz 2004). One way of exemplifying the very wide range of observational tasks defined as visual search is to report the examples given by scholars as to what it constitutes. The following observational behaviours have all been defined as visual search tasks and are referenced at the end of this paragraph. These visual search behaviours are; looking for lost keys, driving, way finding, looking for your socks in the morning, locating tools in a tool box, inspecting aircraft for defects, looking for your car in a car park, a shopper looking for groceries held in a memorised listing, soldiers on patrol and looking for enemy fighters in a crowded market, military target acquisition, searching for lost sailors at sea, predators looking for prey, looking for a friends face in a bar, foraging for edible food in a forest, looking for tumours on an MRI scan, looking for injuries on x-rays, and looking for dangerous objects during security screenings, looking for pens on your desk and watching for pedestrians whilst driving. Biggs & Mitroff, (2013); Cain et al (2013); Cain & Wolfe (2014); Cain et al, (2011); Chan & Chu, (2010); Drury & Watson (2002); Eckstein, (2011) Gramopadhye, et al, (2002);

Palmer, Verghese, & Pavel, (2000); Solman, Hickey & Smilek (2014); Nakayama, & Martini (2011), Wolfe et al, (2016).

In terms of applied or so called "real world" scenarios the following search behaviours have also been defined as visual search tasks; lifeguard surveillance of swimmers in swimming pools, (Lanagan-Leitzel, et al, 2015) security baggage screening by x-ray in airports (Biggs & Mitroff, 2013) and medical image scan diagnostics by physicians (Wolfe et al, 2015).

It has been stated that, as with all search tasks, subsequent behaviour is the result of visual stimuli and cognitive attention decision processes (Palmer, et al, 2000). In this regard visual search is increasingly being understood in terms of memory, decision analysis and reward Nakayama, & Martini, (2011). Visual search can be considered as a function of "representation, attention and understanding" Representation concerns the visual stimuli entering the eye. Attention concerns the filtering of the continually vast amount of stimuli towards those objects of interest. Understanding concerns the cognitive processing of the selected visual stimuli resulting in our object recognition, perception and subsequent behaviour within our visual environment (Palmer et al, 2000; Wolfe & Horowitz, 2004). Furthermore, visual search behaviour is "accomplished by processes and representations common to nearly any visual task" and that "these behaviours are dependent on the representation of visual stimuli, the limits of attention and the integration of information for decision" (Palmer et al, 2000).

Visual search is often not an isolated activity. For instance, driving a car is partly a visual search task as the driver will be looking out for stop signs, speed limits, pedestrians etc (Wolfe et al, 2015). However these scholars complicate their visual search definition by referring to this driving task as "monitoring". The next two sections present the literature on the applied types of visual search and visual inspection definitions which again will show a degree of terminology overlap and interchangeability.

2.2 Visual Inspection Definitions

In terms of discipline jargon, scholars from the cognitive visual psycho-physics background tend to use the term visual search for all observational tasks. In contrast, industrial engineering scholars tend to use the term visual inspection. Therefore the literature presented in this section is largely from the industrial engineering discipline.

When the cognitive component of decision making is added to visual search, many scholars from the industrial engineering discipline refer to this combination as visual inspection (Gallwey 1998(a); Melloy et al, 2000; Nalangula et al, 2006; Rao, 2006; Tetteh et al 2006). It is also reported that a two component representation of visual inspection; search and decide, is the most widely accepted model within industry (See, 2013). The US based Federal Aviation Authority definition of visual inspection is "the process of using the unaided eye alone or in conjunction with various aids, as the sensing mechanism from which judgments may be made about the condition of a unit to be inspected." Drury & Watson, (2002)

Visual inspection has also been described as; "the process of examination and evaluation of systems and components by use of human sensory systems, aided only by mechanical enhancements to sensory input such as magnifiers, dental picks, stethoscopes, and the like" (Moore et al, 2001). It is also pointed out that additional senses and behaviours are used in visual inspection including feel; (moving, pulling twisting, manipulating) sound and smell (Moore et al, 2001; Drury & Watson 2002). Visual inspection is further described as a three step approach of formation, processing and deduction by Gallwey, (1998b). Formation is the learning what to look for search stage. Processing involves the search for objects of interest, a comparison against mental image of what is acceptable and a judgement of acceptability. Deduction involves monitoring the effects of processing.

The observational process involved in conducting a visual inspection has been described as follows; the area to be searched should be viewed as consisting of a series of uniformly sized cells which correspond to the visual lobe or field of view in the eye. Each of these cells, lobes or views can be perceived in a single glimpse or a fixation. A defect is located when the visual inspector fixates on a cell containing the defect (Melloy et al, 2000). Regarding industrial component quality control Gallwey, (1998a) theorises on the number of fixations required for an exhaustive search of a square with sides of 364mm using foveal resolution. This scholar states it will require in the region of 100 fixations using a 10 by 10 row and column pattern.

2.2.1 The Visual Search & Visual Inspection Overlap.

The overlap and interchangeability of visual search and visual inspection reflects the terminology used in the two principal academic disciplines involved; cognitive visual psychology and industrial engineering. Definitions and meanings of visual search and visual inspection are often used interchangeably. Just to exemplify this overlap, the industrial engineering scholars Drury & Watson, (2002) describe visual inspection as; decision making added to visual search. In comparison the visual cognitive psycho-physics scholar Eckstein, 2011; describes the visual inspection of aircraft's fuselage as an example of visual search.

This section will now premise, that any EHS professional who visits a workplace will be conducting a visual inspection as part of his or her duties. The term visual inspection will also be used throughout this thesis to promote readability. Therefore, the following well used terms from the EHS community, (here reproduced from Chapter 1 Introduction) will all be assumed to have involved a visual inspection component; hazard identification, risk assessment, risk audits, risk reviews, risk surveys, safety audits, safety inspections, safety reviews, safety checks, safety tours, safety visits, safety surveys, safety walkthroughs, health and safety visits, health and safety audits, due diligence visits, accident investigations, inspections, surveys, scoping visits, familiarisation exercises, management safety walkabouts and forensic investigations. The advantageous corollary of this argument and the main reason for its consideration is simple; the literature from all visual search scholars can be applied to the findings in this study.

2.2.2 Describing Visual Search Behaviour

Visual search behaviour, or how humans look for defects hazards or targets is an important aspect to understand in the context of this thesis. This is because models of visual search behaviour can explain visual inspection conduct by EHS professionals during workplace inspections. Definitions of the various types of visual search behaviours are now succinctly explained.

2.2.3 Random & Non Random Visual Search

A very basic dichotomy categorises visual search into random and non-random behaviours (Nalangula et al 2006: Nickles et al 2003; Melloy et al, 2000; Wang et al, 1997). Random visual search behaviour is therefore visual inspection conducted with no pattern to the inspector's observations. A non-random visual inspection in contrast, will have a set pattern to its observational conduct. Nalangula et al, (2006) state that better target identification performance should be exhibited by inspectors using systematic as opposed to random visual search strategy. Melloy et al, (2000), state that the only way an exhaustive scan can be carried out, is by using a non- random or in their words "a consistent visual inspection search behaviour, which ensures all areas of the object under analysis are included in the observation". The literature to support consistency in the form of using a set eye scanning pattern, as opposed to random search behaviour is presented in Chapter 3 (Visual Inspection Literature).

2.2.4 Hybrid Search, Foraging Search & Hybrid Foraging Search

There are three further distinct types of visual search behaviour described in the literature; hybrid search, foraging search and hybrid foraging search (Wolfe et al, 2016).

Hybrid search, is the simplest model of visual search behaviour and has been defined as observers searching for "one instance of a target, from multiple types of target held in memory" (Wolfe et al, 2016). An example from lab based visual search inquiry, would include a visual display on a PC screen, set up so that it had one letter L within a group of T letters. Hybrid search would involve participants looking for and finding this letter L from amongst all of the T's.

Foraging search is defined as "observers searching for multiple instances of a single target type or class" (Wolfe et al, 2016). Again from visual search inquiry, multiple letter L's would be displayed within many other letter T's. Foraging search would involve participants looking for and finding as many letter L's from amongst all the T's.

Hybrid foraging search is described as observers searching for "multiple instances of multiple targets (Wolfe et al, 2016). Using the above examples it would be searching for say; all A, B and C letters, from a display that has multiple alphabets included. Hybrid search has been extensively studied in laboratory based studies. Whereas, hybrid foraging search by comparison has been studied with far more applied or "real world" scenarios. (Eckstein, (2011): Nakayama & Martini, (2011). Some scholars argue that many "real world" visual searches exemplify hybrid foraging behaviour (for example see Wolfe et al, 2016). The possible occurrence of hybrid foraging search in the two thesis studies is discussed in detail in section 6.5.3.

2.2.5 The Role of Memory in Visual Search

Memory has an obvious role to play in visual search. This is expressed by a rather axiomatic statement from Cain & Wolfe, (2014) reporting that humans need to remember what they are looking for and remember what was found. Therefore EHS professionals first need to have an attentional set memorised. In short, a list of hazards needs to reside within their memories before they can be potentially recognised during visual inspections. Memory has been described by Baddeley, (2007) as an information processing system composed of a sensory processor with short term memory (also known as working memory) and long term memory components. After receiving sensory data, the short term memory acts as an encoder and retriever between itself and the long term memory. In this way sensory data is encoded and transferred to the long term memory where it can be retrieved for conscious perception.

Scholars have speculated on the role that memory plays during a pre-determined hybrid foraging search. Consider for example; reading a shopping list and seeing the word lemon whilst in a supermarket. Visual working memory serves to retain the object of interest, the lemon, with a conscious perception of its need to be found. On observing a lemon, visual stimuli from this lemon is transferred to the activated long term memory. Here it is rapidly processed by being compared to attentional sets that include lemon like objects, until an understanding is formed of the object and a decision made as to whether that object is indeed a lemon. This results in successfully matching the lemon held in the observer’s memory, with a real one observed in the supermarket’s fruit aisle (Drew et al, 2015; Drew & Wolfe 2013).

How the brain can so quickly processes visual stimuli into to full object recognition objects has also been speculated on. One theory is based on the finding that; object recognition response times in multiple target visual search experiments is a linear function, whereas memory retrieval times for targets is logarithmic (Cain & Wolfe 2014). Here, a possible explanation for this dual time discrepancy is that multiple targets are held within a version of long term memory, and categorised into three processing levels which these scholars categorise into basic entry; a dog or a daffodil, sub-ordinate categories; a short haired poodle or a dwarf white daffodil and superordinate categories; animals and plants (Cain & Wolfe 2014). If these categories were extrapolated to this thesis, then these same processing levels that could possibly apply include basic entry; flammable liquid or cooked food left in the open, subordinate categories; acetone or a meat pie and finally superordinate categories; occupational safety & health hazard or food safety hazard.

The retrieval of categories for further processing by logarithmic elimination rates, is thought to allow for a much faster processing of target recognition (Cain & Wolfe (2014). These scholars further speculate that in “real world” multiple target searches, objects of interest are visually diverse and appear in a variety of colours, shapes and sizes. Therefore, the ability to detect these visually diverse targets, given the linear response time evidence points to the use of logarithmic category selection prior to object recognition.

Evidence about how visual search and memory processes interact during multiple target searches has also been presented by Drew & Wolfe, 2013. These scholars state that both the attentional set (learned targets) and the visual set (what is displayed visually) are used in a hybridised manner during real life inspection scenarios where multiple targets are involved. They further speculate that an x-ray baggage inspector visually searching for knives, guns and explosives can achieve such a task quickly by using a logarithmically structured elimination process. Therefore any observed targets that do not initially resemble memory sets are dropped from attention on successive 50% rates. So the first 50% of sets are dropped then next 50%, then the next 50% and so on. In this way, object recognition during multiple target searches for knives, guns and explosives can take milliseconds instead of many minutes.

2.2.6 What is Known About Current Visual Inspection Custom & Practice

This section will attempt to describe what is meant by current visual inspection custom and practice. This is an important consideration as the research hypothesis will attempt to compare visual inspection performance between systematic visual search users, and their colleagues who were instructed to use their normal visual inspection custom and practice. Furthermore, the experimental design in Chapter 4 (Methodology) calls for ecologically valid conditions which again required control group participants to use their normal visual inspection behaviour. But for such a ubiquitous practice, defining current visual inspection custom and practice remains problematical due to a number of factors, These include idiosyncratic behaviour whereby visual inspection conduct can vary between differing EHS professionals (Hrymak et al, 2015). In addition, visual inspections are normally nested within other requirements of the workplace visit for example the need to produce risk assessments, or conduct safety audits which in themselves vary greatly and thereby affect visual inspection conduct. As mentioned in Chapter 1(Introduction) there is a bewildering range of descriptors, that can affect visual inspection conduct due a lack of accepted meaning. There are also additional hazard identification methods used in conjunction with visual inspection such as asking questions, reading on site documents or checklists (Clift et al, 2011; Hrymak et al, 2015).

Nevertheless, the literature does evidence a fundamental commonality of visual inspection behaviour that can safely be described as, current custom and practice by EHS professionals looking for workplace hazards. This commonality involves EHS professionals walking around the site under analysis, and looking for observable hazards. If these observable hazards are seen, then they will be recorded (Hrymak et al, 2015; Hollis & Bright, 1999; Woodcock, 2014). There are further pieces of evidence to add to this rather obvious finding. The EHS professional is well qualified. They exhibit a minimum time period spent at the site under analysis of at least 30 minutes. They can take further aids or equipment such as torches or mobile phones. Finally they produce findings that are shared with other stakeholders (Drury & Watson, 2002; Hopkins, 2011; Hrymak et al, 2015; Hollis & Bright, 1999; Woodcock, 2014). The use of checklists is more difficult to ascertain regarding current visual inspection custom and practice. A study involving 40 Irish EHS professionals did not find checklists to be featured (Hrymak et al, 2015). However some scholars report their ubiquitous use especially relating to observable, but more technically deterministic hazard identification methods, for example assessing musculo-skeletal risks such as repetitive strain injury amongst call centre workers (Clift et al, 2011) There is also some evidence to show that checklist use may be at its highest, amongst non EHS professionals attempting to discharge their regulatory risk assessment requirements (Clift et al, 2011; Neathey et al, 2006).

Current custom and practice for visual inspection varies due to the wide range of observational tasks required to be conducted. Regarding the aim and research question that this thesis seeks to address, a commonality of current custom and practice for EHS professional will be taken to include; walking around the site under analysis looking for observable hazards, which if seen, are recorded. As will be seen in Chapter 4 (Methodology) this behaviour has been incorporated into the thesis experimental design to produce ecologically valid conditions.

In conclusion, current visual inspection practice is described in the literature by Hrymak et al (2015) and Woodcock, (2014). As stated by these scholars a commonality of all visual inspections is; looking around the area under analysis, recording the hazards observed during the inspection before producing a report for stakeholders. In addition, the lead investigator can attest to this practice as normative and used by him throughout his thirty year career. In addition, as risk assessments are required by law to be written down, the recording of hazards within the Environmental Health and Safety Community is usually by pen and paper and prior to a stakeholder report being produced.

Therefore requiring participants to look for and then record observed hazards by writing them down during the inspection, mirrors Environmental Health and Safety professional practice. Due to the ecological validity requirements of the two thesis studies (see section 4.1.5), this looking around and writing down hazards seen behaviour, was designed into the Kitchen and Aviation studies to reflect current practice.

2.3 The Causes of Visual Search Error

This next section will present a detailed review of visual search error with current theories as to why, visual inspections can often fail to see observable defects, hazards and targets that are present during inspections. In this regard, visual search scholars use the term error literally, to include any situation whereby a defect, hazard or target has been required to be looked for but has not been seen. Error can be used pejoratively. So throughout this thesis, when describing the empirical ability of EHS professionals to observe hazards, the phrase visual inspection performance (VIP) meaning the mean percentage of hazards observed is presented. This section will begin by describing what is known about rates of VIP error, before presenting a taxonomy on why observable hazards are not seen during visual inspections.

2.3.1 The Range of Visual Inspection Performance Error

Here, a brief and general overarching literature and commentary on just how effective humans are at observing defects hazards or targets will be presented. Unsurprisingly there is an expectation of accuracy from lay readers when they consider the efficacy of inspectors engaged in socially important visual search procedures. Accuracy in visual search is explicitly expected within the field of radiology and security screening (Mitroff & Biggs 2014; Wolfe et al, 2005). Similarly the EHS community expects all hazards to be identified during the inspection process (Woodcock, 2014). However when measured in terms of unobserved targets, defect or hazards, VIP is flawed to say the least. It is reported that VIP varies from 9% to 100% (See, (2012). There is evidence that inspectors making visual inspection errors was known as far back as the 1930's. (Nickles et al, 2003 citing Juran, 1935) It is also thought that VIP within the industrial engineering sector is about 80% and in the highly regulated US nuclear power industry VIP is reported to be circa 85%. See, (2015).

As an indication of how error prone visual inspection is can be appreciated by considering how visual search and inspection scholars characterise the practice. Scholars such as Drury, (1999) state that for the aviation industry, "visual inspection is neither easy or fast". Wolfe & Horowitz (2004) simply state that "some search tasks are easy some are not". Drury & Watson (2002) state that errors are not just missed defects, they also include false alarms. Gallwey, (1998b) states that visual inspection problems include the rejection of acceptable objects. He goes on to describe modern industrial inspection practice as; "a sorry picture which belies the somewhat rosy view of some people within it". In a review article, See, (2013) succinctly summarises the literature from the visual inspection discipline by stating that; human inspectors are imperfect, and exhibit large differences in performance across a range of inspection tasks.

So the question now asked is why are humans so unreliable when tasked to look for observable objects that can and do, cause serious consequential problems?. This next section will begin by presenting an overarching categorisation of factors to explain visual search error before detailing a taxonomy of sensory perceptual, cognitive bias and organisational causes for this type of error.

2.3.2 Overarching Factors Affecting Visual Search Accuracy

The fundamental difficulty with visual search is the uncertainty of target location (Cain et al, 2014; Eckstein, 2011; Mitroff et al, 2015; Wolfe et al, 2016). On entering a workplace, observers simply do not know what defects, hazards or targets could be present, what is their frequency and where they are located. These scholars argue that this fundamental uncertainty, gives rise to so many visual search errors. Woodcock (2014) supports this idea by stating that EHS professionals visit workplaces where hazards may be present in a variety of locations, guises and processes. It has been stated that each visual search error is likely to have its own cognitive processes, but that the totality of search error arises from a combination of "attentional, mnemonic, strategic and motivational influences" (Cain et al, (2013). Various scholars characterise visual inspection as cognitively demanding and a difficult task to do well (Biggs & Mitroff, 2013; Drury & Watson 2002; Gallwey, 1998a; Rao et al 2006; Wolfe, 2005; See, 2012).

The construction of categorisations to group the factors that affect visual search and inspection accuracy, varies with the background disciplines involved. A visual cognitive psycho-physics based four category classification, for the factors that limit visual search performance is given by Eckstein, (2011). He states that factors that affect visual search are; degrading resolution away from the fovea, variability in the visual environment, the stochastic nature of visual processing in the brain and finally, limitations in attention and memory. Eckstein, (2011) further states that there are five strategies that the brain uses to optimise visual search and counter these factors that limit performance. These brain strategies are; saliency, target knowledge, reducing uncertainty in target location, optimising the visual system and finally linking reward to target prevalence. A further five category human factors based classification, is here reproduced in Table 2 below, using cited references provided by See, (2012).

Table 2 Factors that Impact Visual Inspection Performance

Task	defect rate, (Fleck & Mitroff, 2007) defect type, (Dalton & Drury, 2004) defect salience, (Geyer et al, 1979) location, (Kane, et al 2009) complexity, (Ainsworth, 1982) standards, (Rao, et al 2006) pacing (Drury & Watson, 2002) multiple inspections, (Dury, et al 1986) overlays, Fox, (1973) automation (Jiang et al, 2007)
Individual	gender, Heidl et al, (2010) age, Wales et al, (2009) visual acuity, (Drury & Watson, 2002) intelligence, (Wang & Drury, 1989) personality, (Thackray, 1993) aptitude, (Drury & Watson, 2002) time in job, and experience, (McCallum et al, 2005) visual lobe, (Gramodpadhye & Madhani, 2001) eye scanning (Nalangula et al (2006); Nickles et al, (2003); Wang et al, (1997) bias (Kane et al, 2009)
Environmental	Lighting, (Drury & Watson, 2002) noise, (Hancock, 1984) temperature, (Woodcock, 2003) time of day, (McCallum 2005) vigilance, (Ghylin et al, 2007) workplace design, (Yeow & Sen 2004)
Organisational	management support, (Wiener, 1975) training and retraining, (Gramodpadhye et al, 2003) instruction and feed forward information, (Drury & Watson, 2002) feedback, (Gramodpadhye et al, 1997) incentives, (Watanappa, 2012)
Social	pressure, (Taylor, 1990) isolation, (Foot & Russon, 1975) consultation and communication (Hillman et al, 1976)

The factors in table 2 above reflect research into particular variables that can affect visual inspection. As will be seen in Chapter 4 Methodology, these variables listed in Table 2 (other than the intervention itself) were not manipulated in order to reflect ecological validity. However, as will be seen in the Chapter 5 Results, the effects from some of the variables in Table 2 including; salience, gender, experience and bias as well as the intervention (noted in Table 2 above as scanning) will be fully discussed.

A taxonomy of sensory perceptual based visual search error, with the categorical headings; scanning errors, recognition errors, decision errors, resource depletion errors, search strategy errors and perceptual set errors is presented by Cain et al, (2013). These scholars explain these as follows; scanning errors occur where the target was simply never seen. In their laboratory based study, Cain et al, (2013) found that nearly half the search errors were due to the targets not being fixated upon. Recognition error occurs when the target was fixated, but not long enough in terms of time to be recognised or evaluated. This factor can include “attentional blink”, whereby the targets fixated within 200-250 milli-seconds of the original are not consciously processed. Decision errors occur when the target was adequately fixated and considered, but not reported as a target. Resource depletion errors occur when the target, in being identified, interferes with subsequent performance by depleting cognitive resources in the searcher. Search strategy errors occur when searchers terminated early, and false errors occurs where a missed target was fixated prior to another target being fixated. Finally perceptual set errors occur when the initial target influences the identification of subsequent targets. Using the example of radiology, a perceptual set error would include; a surgeon looking at an x ray and discovering a broken bone as the first target. Such diagnosticians are more likely to find subsequent broken bones, rather than different lesions, for example cancerous growths. These scholars further state that the visual search misses do not seem to have an underlying cause, instead a number of factors contribute to search misses. Furthermore, attention misdirection, time pressure, anticipatory anxiety, global contextual pressure, working memory resource depletion are all stated as error sources (Cain et al, 2013).

Other scholars list four main causes for task based visual search error as; "target visibility, unknown target set, multiple targets, and low target prevalence" (Biggs & Mitroff, 2014). Target visibility refers to how easy it is to observe the item of interest which includes the following factors; distinctiveness, orientation, and how the target is surrounded. Biggs &

Mitroff, (2014) also support the (unsurprising) idea that high salience targets (their conspicuity, or how much they stand out) are easier to detect than low salience targets. They cite further studies showing that searching for multiple targets, instead of single targets, results in less accuracy, slower response times and impairs attentional guidance.

2.4 Sensory Perceptual Causes of Visual Search Error

The next section of this chapter moves away from a general categorisation of visual search error, to consider specific evidenced phenomena under the three headings of sensory perceptual, cognitive bias and organisational.

2.4.1 Visual Search Error Caused by Memory

Visual working memory has an important role in attentional deployment, but is reduced in capacity during visual search (Drew et al, 2015). These scholars theorise that the process of holding data in the visual search memory reduces the rate of evaluation of new targets (Drew et al, 2015). It is also pointed out that interference between memory sets can also lead to missed targets (Cain & Wolfe, 2014). These scholars further state that categorisation within memory can lead to targets being assigned to specific attentional data sets. Hence a lemon may be held within a “fruit” category in the memory, and a bunch of celery within a vegetable category. Extrapolating these categorisations to “real world” scenarios means that a food inspector looking at a vegetable aisle in a supermarket may well miss the presence of an incorrectly stored lemon (Cain & Wolfe, 2014).

A further possible limitation on visual search performance from memory, is the amount of visual stimuli. In this case, the capacity of visual short term memory is thought to be linked to visual information load and the number of objects in view (Alvarez & Cavanagh, 2004). On top of these issues, there are well known limitations from the human memory system relating to capacity (a limit on what can be memorised) interference (competing versions of memory for the same event) and degredation (a decay in the amount of memory from age or disease) (Baddeley 2007). The number of items that can be held in the working memory is popularly characterised as “Millers magic number seven, plus or minus two” (Miller, 1956). These memory related causes of error are considered to be specific to all of us, and therefore will be acting upon all EHS professionals. Cain et al, (2013) speculate that once visual searchers consider that multiple targets exist, memory itself is constrained or lessened as if it is being pre-allocated for dealing with targets. Hence, there is less working memory available to deal with subsequent targets. Therefore, remembering the location and identity of a target is thought to consume working memory and thereby diminish cognitive resources.

2.4.2 Subsequent Search Misses

The phenomenon of subsequent search misses, whereby the observation of a target negatively affects the observation of subsequent targets has been identified from laboratory experiments. An example would be an airport security x-ray baggage screener finding an initial target such as a bottle of water in a suitcase, which somehow lessens the observation of subsequent and harder to observe targets such as knives in the same suitcase (Fleck et al, 2010). Another study describes how radiologists are more likely to miss additional abnormalities if they first find a fracture on an x-ray scan (Mitroff et al, 2015). These scholars also describe how pathologists are more likely to miss additional cancerous lesions on microscope slides after they find the first anomalous lesion.

Some scholars state that multiple target search is less accurate than single target search which may be down to the decision as to when to stop searching. The reason for this may be due to “attentional disruptions” and the “depletion of cognitive resources” subsequent to first target detection Clark et al, (2014). These scholars further state that subsequent search misses is a generalisable concept and not confined to laboratory experiments. Finally these scholars state that target crowding can also lead to error when compared to evenly dispersed targets, which may be due to differences in serial and parallel cognitive processing of visual stimuli.

There are two principal theories to explain the subsequent search misses phenomena during multiple target visual search; the perceptual set theory and the resource depletion theory (Mitroff et al, 2015). The perceptual set theory states that the first target observed creates a bias, whereby the observer is subsequently more likely to look for, and therefore find, additional targets that match the visual characteristics of the first target. For example, a radiologist on observing a fracture on an x-ray chart is more likely to observe further fractures and less likely to observe other non-fracture type lesions such as inflammation. In agreement with this perceptual set theory Wolfe et al, (2016) offers support by explaining that attention is guided to items with similar features to targets that have already been observed. Furthermore, multiple target visual search error can occur if observers are conducting visual searches using hybrid foraging behaviour. These scholars state that not all targets will be observed due to a cognitive characteristic predicted by hybrid foraging behaviour. This results in a trade off between time spent observing with the number of targets found. This effect is further detailed in section 2.4.7.

The resource depletion theory states that observing the first target creates a cognitive burden as the observer must hold this target and its location in their working memory. This could also be considered as a re-fixation distraction issue, a possible visual search error suggesting that short term memory is not capable of distinguishing previously searched areas (Clark et al, 2014). It is further speculated that these types of multiple target visual search error could amount to circa 15% of all visual search errors. Despite this 15% figure being derived from controlled laboratory conditions, Mitroff et al (2015) argue that there is currently no reason why this error rate cannot be applied to "real world" searches. These scholars also speculate that situational factors can exacerbate subsequent search misses which include; the search context, time pressure and finally expectations regarding the number of targets present. In support of this idea, Cain et al, (2011) report that stress in the form of anticipatory anxiety can also detrimentally affect multi-target search accuracy.

The context of the search can also affect subsequent search misses in terms of the importance of the target (Biggs & Mitroff, 2013). This is supported by a study that used laboratory conditions to compare professional baggage searchers with non-professional searchers (University undergraduates) and that also manipulated the variable of target importance. They found that non-professionals were more likely to miss a second target when designated dangerous and professionals were more likely to miss a second target when designated safe. These scholars conclude that reducing subsequent search misses can be achieved by categorising all targets as dangerous (Biggs & Mitroff, 2013). A similar finding is reported in a study by Wolfe et al, (2017) who also manipulated target importance during a lab based search and found the greater observation of more importantly designated targets.

Search satisfaction, or when to stop searching, is also thought to be linked to previous experience in that searchers quit when they think they have found enough targets or non are detected for a time (Cain et al, 2014). These scholars also state that factors contributing to subsequent search misses include; misdirection of attention, time pressure, global contextual pressure, and attentional refractory period.

Subsequent search misses, is therefore a well established phenomenon, whereby “finding one target can cause a searcher to miss additional targets” (Mitroff & Biggs, (2014b). It has also been described as “success breeding failure” (Mitroff et al, (2015). The scholars; Clark et al, (2014) also state that searching for a second target is cognitively challenging and is sensitive to contextual influences. These scholars further suggest that satisfaction of search problems may be addressed by fixed duration time searches rather that fixed object searches. Wolfe, (2013) states that search termination behaviour is important when observers do not know the number of targets present, which will be the case in many “real world” search scenarios.

2.4.3 Target Prevalence

The prevalence (or infrequent appearance) of targets within a search area, has been repeatedly shown to affect search accuracy. The literature demonstrates that the less likely the target is to appear, the less likely it is to be observed during searches. Therefore visual search behaviour that is successful at a high prevalence rate may not be so accurate in “real world” searches with much lower prevalence rates (Mitroff & Biggs, 2013; Mitroff & Biggs, 2014);

These scholars report their concern in a study where they manipulated target prevalence in baggage screening so that it was between 0.078% and 3.72% of luggage passing through x-ray machines. Under these prevalence rates, low levels of observation were reported with near 90% error rates (Mitroff & Biggs, 2013). This low level of detection raises concerns as many “real word” searches, including radiology and security screening are conducted under ultra- rare target incidence rates. A further study, conducted using a smartphone app called “Airport Scanner”, simulated x-ray baggage inspection” (Mitroff & Biggs, 2014). They found that for ultra rare items (less than 0.05% of luggage containing targets), the error rate was about 45%. They describe this error rate as “disturbingly poor.” For low prevalence targets, scholars have suggested planting additional targets to maintain attentiveness and prevent inappropriate low thresholds of search quitting behaviour (Biggs & Mitroff, 2013). In contrast visually searching at the other end of the scale whereby target prevalence is high, is also associated with detrimental VIP (Gallwey and Drury, 1996). These scholars argue that task complexity is a difficult obstacle to overcome and speculate that searching for more that about 6 defects, becomes difficult.

2.4.4 The Speed Accuracy Trade Off

The cognitive phenomenon of the “speed accuracy trade off ” has been reported by Biggs & Mitroff, (2014b) and Melloy et al, (2000). Basically, a faster visual search can be performed, but it will be accompanied by a reduced level of accuracy. These scholars found that professional searchers were more accurate when they were slower in their visual inspections. It is also reported that visual search tasks are characterised by speed and accuracy, but these two variables are inversely related with accuracy decreasing with increasing speed (Melloy et al, 2000). The same inverse relationship has been reported in a study of baggage security staff by Wales et al. (2005).

In a study of ten surveyors tasked with producing house condition reports, only one surveyor found all relevant housing defects. This low VIP rate was also attributed to the speed accuracy trade off (Hollis & Bright, 1999). They found that surveyors who spent more time inspecting, observed and recorded more housing defects. Some scholars state that nothing has yet been found that eliminates search error when the searcher is under any type of time pressure (Cain et al, 2013). Decreased visual search accuracy has also been found when time pressure has been increased in lab experiments (Fleck et al, (2010). Finally, the balance between speed and accuracy has been described as having safety, dependability and cost implications if not appropriately controlled (Cain et al, 2013).

2.4.5 Diligence &Vigilance Effects

Human variability amongst inspectors in terms of search accuracy, is well reported (for example see Drury & Watson, 2002; Gallwey, 1998; Jiang et al 2000; See, 2013; Solman, (2014). These scholars all state that a lack of diligence, which they describe as a motivational factor, is a cause of visual search error. It is further reported by Solman et al, (2014) that diligence is a factor which is distinct from vigilance decrement and in-attentional blindness.

Vigilance is defined as the “ability to focus attention and detect signals for an extended period of time, under conditions where signals are intermittent, unpredictable and infrequent (Lanegan-Leitzel et al, 2015). These scholars report that lifeguards are given breaks or rotate positions every 20-30 minutes to minimise visual search decrement. This decrement is thought to be caused by cognitive resource depletion and signal sensitivity decreases.

A “vigilance decrement” phenomenon, has been known for half a century (Drury & Watson, 2002). These scholars explain that the probability of observing defects decreases with time.

They also state this decrement is particularly noticeable after the first 20-30 minutes of a visual inspection task. However, the data for this phenomenon is lab based, and these scholars point out the uncertainty of generalising to “real world” search tasks. Other scholars such as Gallwey, (1998b) further cast doubt on vigilance decrement stating it remains an open question.

2.4.6 In-Attentional Blindness

In-attentional blindness is explained by Aimola-Davies et al, (2013) as; “once an attentional set is adopted, an unexpected object or event may go undetected if it does not share the same set properties”. This was shown in the well known example; “The Invisible Gorilla”, a study by Simons & Chabris, (1999) whereby participants were asked to count the passes between basketball players on a video display. It was found that 42% of participants did not notice a gorilla suited actor walking through the basketball players. In-attentional blindness can also occur is expert observers (Drew et al, 2013). These scholars detailed in-attentional blindness in a group of 24 radiologists, who were selected to observe MRI scans of potentially tumerous lung tissue on PC screens. MRI’s (magnetic resonance imaging) together with PET (positron emission tomography) and CAT scans (computed tomography) are routinely used diagnostic visual search aids. In this study, certain MRI scans were altered so that the image of a gorilla, about the same size of a match book, was blended into particular scans and designed to stay visible. The results were that 83% of the radiologists did not “see” the gorilla on the scans. This was despite eye tracking technology showing that the majority of radiologists had looked directly at the gorilla image. The authors of this study, clearly state that these findings are not an indictment of radiologists. They state that the radiologists would have identified the gorilla far better, if they were expecting such an image. Instead this experiment serves to present more evidence of human limitations in attention and perception capability, when engaged in visually demanding search tasks. A related phenomenon; change blindness, is also reported in the literature which consists of observers “failing to notice large and significant changes in scenes that they are scrutinising” Lanagan-Leitzel, et al, (2015).

2.4.7 Hybrid Foraging Search Error

Those visual search scholars interested in “real world” hybrid foraging search behaviour (looking for “multiple instances of multiple targets) have extrapolated “marginal value theorem” into their inquiry. This theorem was first published in 1976, and models animal search and feeding behaviour (Charnov 1976). This theorem is of interest to visual search scholars as it examines how much time an animal will spend looking for a food target in a patch of land, before stopping their visual search and moving to the next patch of land to resume their search there. It throws light onto human visual search behaviour as it demonstrates that the resultant effect of marginal value theorem is a demonstrated detrimental effect on VIP. This deleterious effect is achieved by cognitive processes fundamentally limiting the observation of targets to match previous experience.

To explain, the marginal value theorem posits that the average time spent by an animal searching for food is based on previous successful search times. An example to consider is a gibbon, searching for an apple in a tree. If it historically takes five minutes for that gibbon to find an apple, then it will stop searching any new tree after five minutes, regardless of the actual number of apples there may be in the tree it is searching. In terms of a gibbon finding all the apples in one tree, marginal value predicts that this type foraging search behaviour will not be particularly successful. This is because the accurate observation of all apples in a tree is not related to the number of apples in that tree, but rather the time normally taken to successfully observe an apple. Therefore hybrid foraging search behaviour in humans, if it follows the marginal value theorem, is fundamentally limited when it comes to successfully observing all targets required (Wolfe et al 2016). Hybrid foraging search behaviour will be successful however, when it comes to quickly finding easily seen, or in the words of these scholars, “low hanging fruit”.

Visual search scholars have demonstrated hybrid foraging search behaviour in humans but only under laboratory conditions whereby participants are tasked with finding targets on a PC screen. (Cain et al, 2013; Fougnie et al, 2015; Wolfe, 2013; Wolfe et al, 2016; Zhang et al, 2015). These scholars state that the marginal value theorem effects were demonstrated to an extent in these studies. If these scholars are correct in generalising these lab findings to the “real world” then the same limiting effect on human visual search performance from the marginal value theorem should also apply. In effect, if humans visually search using hybrid foraging behaviour the point at which they decide to stop searching will be a function of previously successful search times, and not the number of defects hazards or targets.

Some scholars state that human search heuristics are geared towards a “hybrid foraging” model which rewards rapid progress with reward. Reduced reward intervals interferes with this process and requires a more diligent approach to overcome, which goes against our human internal biases (Cain et al, 2014). These scholars summarise this point by stating that if we humans don't find our object of interest quickly, we tend to stop looking and move on where the pickings may be better.

Again, speculatively applying this theoretical effect to the "real world" some scholars report that where an exhaustive search with high visual search performance is a necessity, hybrid foraging search behaviour could be highly problematic (Cain et al, 2014; Wolfe, 2013; Zhang, et al 2015). This is because hybrid foraging visual searchers rarely continue observing until all targets are identified. Instead, people stop searching and move when target identification falls below average rates of observation as described by the marginal value theorem (Cain et al, 2014; Wolfe, 2013; Zhang, et al 2015).

Therefore applying hybrid foraging visual search behaviour in such areas as; medical image diagnostics or baggage security, will limit the number of targets found (Cain et al, 2014; Wolfe, 2013). This effect is clearly unwanted when the consequences are high such as missing a bomb in luggage being conveyed to a passenger aircraft (Cain et al, 2014).

Extrapolating this effect to EHS professionals, could mean that the decision by visual inspectors to stop searching a workplace, will not be a function of the number of hazards in that workplaces. But it may be a function of how many hazards were observed in previous workplace visual inspections. Therefore, their visual search performance will not be exhaustive and all observable hazards will not be seen. This effect is also redolent of the availability heuristic (Kahneman, 2011) and some visual search scholars summarise this hysteresis effect by stating that recent experience "changes the estimate of the current state of the world" (Zhang et al, 2016)

However, it currently remains that the effect of marginal value theorem on hybrid foraging search behaviour is still open to debate. Some scholars state that marginal value theorem describes human hybrid foraging search behaviour very well (Wolfe et al, 2018). But some scholars state that the recent experience of the visual inspector is more likely to influence future foraging behaviour than the average as predicted by marginal value theorem (Zhang, et al 2015; Zhang et al, 2016). In contrast, these scholars found that within a lab experiment setting, human hybrid foraging search behaviour is better predicted by recent search event performance rather than marginal value theorem. Other scholars report that lab experiments showed interactions between decision making and that visual search, does not necessarily follow the hybrid foraging search behaviour (Fougnie et al 2015). Other scholars report that assigning importance criteria to targets can affect hybrid foraging search behaviour resulting in a higher number of important targets found (Wolfe et al, 2017). In summary, there are various scholars who remain of the opinion that in "real world" human visual search behaviour, is best explained by hybrid foraging search behaviour (for example see Wolfe et al, 2016). However this position could very quickly change, as other scholars have conducted studies whereby the marginal value theorem has not fully modelled human hybrid foraging search (see for example Fougnie et al 2015; Zhang, et al 2015; Zhang et al, 2016).

2.5 Organisational Causes of Visual Search Error

This section discusses the problem of ambiguity in relevant professional standards and includes a critique on how visual inspection conduct is controlled (or not).

2.5.1 The Inherent Ambiguity within Definitions & Standards

As detailed in Chapter 1 (Introduction) there is currently very little clear guidance as to how to look for workplace hazards. This ambiguity is not confined to visual inspection. Aven, (2011) as well as Johansen & Rousand, (2015) point out that ambiguity is widespread within the EHS community with the concept of risk assessment itself, subject to differing interpretations.

Ambiguous descriptions can be a source or error in terms of inspectors not correctly classifying an object of interest as a hazard. In "real world" scenarios, multiple target searches will include objects of interest that are visually diverse and can appear in a variety of colours shapes and sizes (Cain & Wolfe, 2014). Therefore targets that are described ambiguously can lead to error. For example, consider a food inspector at a customs check point who has been directed to seize all fish and cheese due to a pathogenic organism contaminating such edible items. In this example, Cain & Wolfe, (2014) pose the question; what is the probability of successfully observing caviar covered crackers and a raspberry topped cheese cake that are stored in a picnic basket?.

Regarding industrial inspection, objects of interest are often difficult to define and measure. Standards are many but consensus definition on what is an acceptable defect or not, is difficult to achieve (Gallwey, 1998a). It is also reported that confusion occurs when inspectors are directed to inspect, but given vague information about the level of inspection required (Drury & Watson 2002). Such vagueness they speculate, translates into a question of how much time is available for the inspection as the more detailed the inspection, the more time is needed. These scholars further report that inspectors see their roles as being there to inspect, and not to inspect quickly. The result of this vagueness, means visual inspectors will try to interpret differences between a general or detailed inspection, into how much time they should spend on the inspection task (Drury & Watson, 2002).

This idea that ambiguity causes confusion is supported by Forster & Douglas, (2010). These scholars argue that that any building condition visual inspection tasks are inevitably subjective due to the interpretation required. Furthermore visual inspections surveys conducted for building condition purposes, are again subjective judgements, that cannot but give rise to surveyor variability (Kempton et al, (2000).

Even though visual inspection plays such an important task, many scholars have stated their dissatisfaction with the efficacy of the risk assessment and safety auditing process in general. It is stated rather axiomatically that “an unidentified hazard, is difficult to defend against” Aven, (2011). Similarly it is stated that it is of paramount importance that for the prevention of accidents ill health and their consequences, that all hazard are identified (Pinto et al, 2012). Furthermore, knowing “the hazard profile of an organisation is a necessity and compliance auditing is of little value if this identification phase is not conducted appropriately” (Makin and Winder, 2008). However, it has been stated that the identification of workplace risks is not easy and that the use of simplistic risk assessment tools and basic training is flawed (Clift et al, 2011). Similarly risk assessments should identify all workplace hazards, but this is difficult to do in practice (Pinto et al, 2013). Referring to construction safety, it has been reported that poor hazard identification by employees, safety managers and inspectors is a serious concern (Albert et al, 2017; Dzeng et al, 2016).

Some scholars have little confidence in current safety auditing process. They argue that results from the auditing of occupational safety and health management systems should be treated with caution (Blewett & O’Keefe 2011). These scholars list nine flaws in safety auditing namely; unintentional auditor error, deliberate auditor fraud, inappropriate client influence on the auditor, failure to allow worker participation, excess paperwork, inaccurate audit scoring; confusing audit criteria and finally a lack of auditor skill and independence. As a further example; of five selected occupational health and safety audits conducted in

Australian workplaces, 60% did not have sufficient hazard and risk identification components in the documentation analysed (Robson et al, 2010).

In summary, these negative sentiments are succinctly encapsulated by Wadsworth & Walters, (2014). They report the concerns of many interested in professional safety conduct by stating that; the lack of qualification and competency criteria allows little confidence, that many of the consultants offering health and safety services, are actually competent to do so.

2.5.2 The Professional Control of Visual Inspection Conduct

This section will present a critique on the overall control of visual inspection conduct. It will discuss the efficacy of societal control over visual inspection in Ireland and the UK, which is exercised primarily through EHS professional practice guidance (discussed in Chapter 1 introduction), the legal concept of reasonable foreseeability and the tort of negligence. The "reasonably foreseeable" doctrine, is a legally applicable concept regarding the standard of hazard identification required by professionals (White, 1993; Tomkins, 2010). Under the Safety Health and Welfare at Work act 2005, Irish employers are expected to identify all reasonably foreseeable workplace risks. Risk assessment is regarded as the principal method of discharging this duty (HSA, 2015). An example of what is reasonably foreseeable, is presented by White, (1993). He reported on a UK civil case where a paraffin heater was left on the pavement of a street, undergoing repairs. The heater was knocked over causing burns to a pedestrian. White, (1993) reports that in Hughes v Lord Advocate (1963) 1 All E.R.75 (H.L.(Sc.), that burns due to inappropriate paraffin storage was "damage, a reasonable man should have foreseen". In effect the paraffin heater should have been seen and the risk rectified by appropriate storage. This gives some indication as to what should be identified by a competent risk assessor and from this judicial case; a flammable liquid in open view.

Although an overarching legally based professional negligence standard is expected from inspectors, there are few if any, empirical standards by which to judge the VIP of inspectors other than retrospectively when a court case occurs. However, it is reported that those EHS professionals who rely on the "we have always done it this way" doctrine will find this strategy flawed. Two legal cases from Ireland; O'Donovan v Cork County Council [1967] IR

173 and Roche v. Peilow [1986] ILRM 189 make the clear point that custom and practice is no defence in the tort of negligence. In effect, claiming that this is how risk assessments or safety audits (with their nested visual inspections) have always been done, will not be a sufficient defence if those same inspections resulted in observable hazards being missed and injury or financial loss results. In this regard, it is recommended that some form of recording the thoroughness of the inspection is needed, simply for the inspector's own protection (Woodcock 2014).

2.6 Cognitive Bias Causes of Visual Search Error

In conducting visual inspections, EHS professionals need to make decisions about the defects, hazards and targets they observe. Psychologists have long shown that human decision making can be rational and logical, but can also be the opposite. Decision making can also result in conclusions that are based on heuristics, are simplistic, stereotypical, biased and therefore flawed (Aronson, 2013). There are many examples of heuristics (mental short cuts) and bias but a commonality is that our expectations "blur" our vision. What we see first and what we've seen last affects our overall judgement (Fiske & Taylor, 1991; Gilovitch et al, 2013). The influential "cognitive miser" principle states that; human decision making can be rational and logical, but can also be inaccurate if the easiest and quickest decision making processes are used, rather than the more cognitively demanding thought processes (Fiske & Taylor, 1991).

As decision making is inherent in visual inspection, it should not be a surprise to find heuristics and bias influencing visual inspection outcomes. Daniel Kahneman, a Nobel prize winner and described by Professor Steven Pinker from Harvard University as the "worlds most influential living psychologist" has long investigated this conceptual area, often referred to as cognitive bias. In his international best selling paperback, Thinking Fast and Slow (Kahneman D, 2011). He considers intuition to be a form of object recognition and discusses how this intuition can easily be deceived in experts and lay people alike. Kahneman, (2011) further argues that experts need to constantly verify their decisions, and make their own internal checks to ensure they are actually correct, rather than relying on their intuition. An example he gives is confirmation bias. This occurs when experts believe what they already think or expect, without checking the evidence. It results in experts and lay people alike finding it difficult to disbelieve their own deeply held pre-conceptions. For example a high priced bottle or wine must be better than a cheap one. When this cost based theory is not tested or checked, errors can occur as price does not always guarantee quality. Applied to workplace inspection scenarios, EHS professionals may consider a workplace safe from its superficial appearance rather than investigate appropriately for evidence of safety.

Cognitive and motivational bias as applied to the EHS community, has been commented on by Montibeller & Winterfeldt, (2015). They report their surprise, that bias as applied to the field of risk analysis, has not has not yet been subjected to any relevant research. However in one study (Kempton et al, 2002) confirmation bias in building surveyors was reported. These scholars demonstrated that surveyors who were shown dilapidated exterior views of buildings, predicted higher levels of internal disrepair. In another study, Kempton, et al (2002a) also claims to have seen evidence of the anchoring heuristic amongst undergraduate surveyors when presented with initial data on the building lifetime expectations. Referring to visual inspections conducted by housing surveyors, Kempton, (2004) again states that their results will be influenced by heuristics and bias.

Outcome bias, is described as a tendency to ignore warnings (New Scientist, 2015). This was exemplified as the principal factor in the Columbia space shuttle disaster in 2003. Here, prior examples of insulation detaching from the shuttle were routinely noted. However, as they caused no discernible effects they were not regarded as presenting a risk and were effectively ignored. It is advised that to counter outcome bias, “negative details” should always be considered, which in itself will bias the decision maker towards considering adverse consequences prior to them occurring (New Scientist, 2015). In a generalised analysis of professional judgment Gilovitch et al, (2013) point out that the mental processes that we all possess can lead to both accurate as well as dangerously flawed judgement. Furthermore, Kahneman & Klein, (2009) state that the intuitive judgement of some professionals can be impressively skilled, but remarkably flawed in others. To conclude this analysis of the causes of visual search and inspection error, a summary listing can now be presented as follow in Table 3 below;

Table 3 Summary List of Visual Search Error

Cause	Brief Explanation	Reference
Limitations in Memory	Holding data in visual search memory reduces the rate of evaluation of new targets	Drew et al, (2015)
Interference between memory sets	Missing that a lemon is incorrectly stored in the vegetable section	Cain & Wolfe, (2014) Baddeley, (2007)
Capacity of memory	Too many objects to memorise correctly	Alvarez & Cavanagh, (2004); Baddeley, (2007)
Memory degredation	A decay in the amount of memory from age or disease	Baddeley, (2007)
Subsequent search misses	The observation of a target negatively affects the observation of subsequent targets	Clark et al, (2014); Fleck et al, (2010); Mitroff et al (2015).
Target prevalence	Lower VIP is related to very low levels of target prevalence	Mitroff & Biggs, (2013) Mitroff & Biggs, (2014) Solman et al, (2014)
Speed accuracy trade off	The speed of search and observational accuracy are inversely related	Biggs & Mitroff, (2014b) Melloy et al, (2000) Hollis & Bright, (1999)
Vigilance and Diligence	How well the observational task is conducted, affects VIP	Drury & Watson, (2002) Gallwey, (1998a)
In-attentional blindness	When a given attentional set is adopted, an unexpected object may go undetected if it does not share the same set properties	Aimola-Davies et al, (2013); Drew et al, (2013)
Hybrid foraging search behaviour	VIP is influenced by previous mean times taken to observe targets	Cain et al, (2013) Charnov, (1976); Wolfe et al (2016)
Expert judgement	The performance of professional judgement varies	Aronson, (2013) Gilovitch et al, (2013) Kahneman, (2011)
Confirmation bias	A tendency to see what you expect to see	Gilovitch et al, (2013) Kahneman, (2011)
Outcome bias	A tendency to ignore warnings	Gilovitch et al, (2013) Kahneman, (2011)
Ambiguity in definitions	A lack of precision in defining a defect hazard or target allowing that object to be mis-interpreted as not of interest	Aven, (2011) Johansen & Rousand, (2015) Cain & Wolfe (2014)
The lack of guidance in visual inspection conduct	Idiosyncratic behaviour resulting in some objects of interest not being observed due to visual inspection conduct	Drury & Watson, (2002) Cain & Wolfe, (2014) Clift et al, (2011)

Many scholars have long pointed out just how difficult and error prone visual search is in practice. However, even when legally required, there is still no consensus as to how exactly visual inspections, risk assessments and safety audits are to be conducted. This has resulted in variability of conduct and has been reported by Aven, (2011); Hrymak et al, (2015); Lenhardt & Beck (2016) and Johansen & Rousand, (2015). The literature in this field has empirically reported the limitations of EHS professionals in not observing, readily observable workplace hazards during their inspections. These limitations stem from sensory perceptual, cognitive bias and organisational phenomena. Having reviewed the causes of visual search error, the next section addresses what can be done to improve the situation.

2.7 Improving Visual Inspection Performance

What emerges from the literature are ongoing efforts into how VIP can be improved. There are three main approaches noted in this literature; proceduralising conduct, training and using visual search consistency. To begin with, there are examples where the proceduralisation of visual inspection conduct has been attempted and required search behaviours documented. These examples are; aircraft maintenance (Drury & Watson, 2002) bridge inspection (Moore et al 2001) industrial component inspection (Gallwey, 1998a) the Chartered Surveyor method for building inspections (Hollis, 2000) medical physical examinations (Walker et al, 1990) and finally two very concise examples from lifeguarding, (Lanagan-Leitzel, et al (2015) and fairground ride equipment (Woodcock, 2013). What emerges from sections 2.7.1 through to 2.7.5 are specific attempts to proceduralise visual inspection. These are the only published studies that describe required visual inspection behaviours. They are included here for reference and for comparison to the proposals made for visual inspection in Chapter 7 (Conclusions and Recommendations).

2.7.1 Aircraft Maintenance Visual Inspection Procedures

In the aviation maintenance sector, there are four defined levels of visual inspection, each with increasing complexity (Drury & Watson 2002).

- Level 1; Walkabout; a general condition check.
- Level 2; General; an exterior inspection to detect irregularities
- Level 3; Detailed; an intensive visual examination with the use of all aids and access

- Level 4; Special Detailed; An intensive time consuming and highly technical examination using level three methods above, but including invasive techniques such as x-ray scanning, fluid injection, stress mechanics and telemetry

Furthermore, the actual procedures for the visual inspection of aircraft and their components has been specified. The steps involved here are summarised by Drury & Watson, (2002) who present a range of human factors based practices designed to maximise VIP. They begin by offering the following advice regarding how a visual inspection should be conducted. This advice is here summarised as; general looking is not enough, as the inspector, you should continually ask; what is wrong with this picture?, be inquisitive and question whether you have seen this before. Move, shake, pull, twist, and push all parts possible. A summary of their recommendations for the entire visual inspection process is detailed in the Table 4 below (after Drury & Watson 2002).

Table 4 Visual Inspection Procedures for Aviation Maintenance Technicians

Function	Visual Inspection Description
1. Initiate	All processes up to accessing the component. Get and read work cards. Assemble and calibrate required equipment.
2. Access	Locate and access inspection area. Be able to see the area to be inspected at a close enough level to ensure reliable detection.
3. Search	Move field of view across component to ensure adequate coverage. Carefully scan field of view using a good strategy. Stop search if an indication is found.
4. Decision	Identify indication type. Compare indication to standards and work cards for that indication type.
5. Response	If indication confirmed, then record location and details. Complete paperwork procedures. Remove equipment and other job aids from work area and return to storage. If indication not confirmed, continue search.

2.7.2 Bridge Maintenance Visual Inspection Procedures

Moore et al, (2001) report on five defined levels or types of visual inspection that can be carried out by US Government appointed bridge inspectors. These are as follows;

- Initial; first inspections
- Routine; which are scheduled regularly to assess safety,

- In depth; described as close up and hands on,
- Damage; to evaluate any human or environmental challenge
- Special; to monitor a specific defect or condition. Such as strain gauges and potentiometers to monitor any movement creep or fatigue.

2.7.3 Industrial Component Visual Inspection Procedures

Gallwey, (1996b) provides a visual inspection proceduralisation via his published training curricula for industrial inspection. This curricula, is here summarised as follows;

- Items inspected exactly as in situ.
- Trainees actively participate.
- Feedback always given to trainees by more experienced inspectors.
- Mistakes corrected immediately.
- Level of difficulty progressively raised.
- Inspectors are required to practice.

2.7.4 The Chartered Surveyor Method of Visual Inspection

Visual inspections are extensively used by members of the Royal Institute of Chartered Surveyors, (Hollis, 2000; Reddin, 2016; RICS, 2002; RICS, 2010). Chartered surveyors routinely visually inspect (survey in their jargon) buildings and their curtilages for construction defects such as damp, wood rot, dis-repair etc. This Institute refers to visual inspection as a "survey" but it is essentially the same thing as a visual inspection; entailing a visual search strategy designed to systematically observe all areas of the area under analysis which is normally a building, or some part thereof. Their proceduralisation of conduct begins with a requirement of the surveyor to select one constructional element from a room or area at a time, such as a ceiling or wall, and to thoroughly observe that element before moving to the next element. A common order used to select the element to observe is; ceiling, walls, floor, windows, chimney breast and finally doors.

2.7.5 Lifeguarding & Fairground Ride Equipment

Although the proceduralisation for lifeguarding and fairground visual inspection conduct given here is very concise, it does at least point to forms of visual search guidance. So for consistency of reporting, it is included here. Lifeguards are recommended to "scan from point to point, rapidly watching all patrons in the area". (Lanagan-Leitzel, et al (2015) citing the American Red Cross, (2007). Regarding EHS professionals who inspect fairground ride equipment Woodcock, (2014) reports an informal visual search strategy described as “starting at the outside and moving into the centre”.

2.7.6 Medical Physical Examinations

The medical profession has long proceduralised a visual search strategy used during the physical examination of patients for diagnostic reasons (Walker et al, 1990). Medical practitioners are recommended to conduct a visual search or in their jargon, a physical examination of a patient in the following summarised order; head, eyes, ears, nose, mouth, face, neck, anterior torso, posterior torso, thorax, abdomen, and finally proximal extremities.

2.7.7 Training Based Methods

By far the greatest amount of literature relating to improving visual inspection conduct involves training which is here taken to include practice and feedback. Training (the phrase “attentional set” improvements is sometimes used by visual search psychologists) has been described as being the most amenable, influential and practical intervention for improving visual search. (Biggs & Mitroff, 2013; Eckstein, 2011; Gallwey, 1998a; Gramopadhye, et al, 2002; Wolfe, 2005 and Wolfe et al, 2015). In an overarching comment Eckstein, (2011) states that in “real world” visual search professional practice the number of defects, hazards or targets is unknown, but the situational environment requiring the search, cannot be changed. Therefore the search accuracy of the observer is “arguably the weakest link” and therefore offers the greatest potential for improvement. (Eckstein, 2011). Training has also been found beneficial for visual inspection speed, accuracy and in resolving ambiguities present in the inspection process (Gramopadhye, et al, 1997; Gramopadhye et al, 1997b; Wang et al, 1997; Swaninger, 2005; Woodcock 2014).

In two studies; Albert et al, (2014); Albert et al, (2017) it is reported that a ten term mnemonic based training method for construction site workers, improved VIP. Their 2017 study demonstrated very high levels of VIP improvement to over 70% (see Table 7). Feedback for inspectors on their search performance is also routinely commented on as a beneficial factor in visual search. Examples include Gramopadhye et al, (1997b) and Gallwey, (1998a) who state that feedback is beneficial due to inspectors being able to “pay attention” to their inspection strategy.

In modern educational practice, the PC is ubiquitous and so, it is no surprise to find that software for visual search training packages have been developed, tested and beneficial results for VIP reported. Gramopadhye et al (2000) report on a computer based training programme known as “ASSIST” which was specifically designed to meet visual inspection training for aircraft maintenance technicians. This computer based training programme has been adopted by the US based Federal Aviation Authority for training for aircraft maintenance technicians (Gramopadhye et al, (2000). Sadasivan, (2009) reports that a refinement of the “ASSIST” computer aided training programme, “GAITS” (General Aviation Inspection Training System) is set to become general aviation industry standard for inspection training. It is also reported that virtual reality was better received for training in visual inspection when compared to using PC’s by Vora et al (2002).

In summary, studies that have evaluated the effects of training on VIP (including feedback and practice) have featured highly in the visual search literature. However, which training pedagogies and curricula are recommended for workplace inspections is a point that will be looked at in Chapter 7 (Conclusions and Recommendations).

2.7.8 Using Consistency in Visual Search

This section will detail the published studies on the issue of consistency in visual search. Consistency is explained here as; repetitively using the same visual search behaviour for different observational tasks. For example a consistent observer may always begin a visual search at the top of the object of interest and work their way down. Other consistent observers may start at the bottom and vice-versa.

Unfortunately, there are visual search studies that do not detail the type of visual search behaviour other than describe self-reported behaviour as consistent. Such studies are presented in this section. In contrast, there are three studies that specify the exact nature of this consistency by specifying the eye scanning patters used by observers. These studies are; Nalangula et al 2006; Nickles et al, 2003 and Wang et al, 1997. These three studies are presented in detail in Chapter 3 (Visual Inspection Literature).

The literature generally supports the idea of consistency as an approach to improve VIP (Biggs & Mitroff (2014). For example it is stated that unknown target set error can be mitigated by a consistent search strategy which can lessen cognitive depletion. This because the searcher's short term memory is not used to remember what areas have, or have not, been searched (Biggs & Mitroff 2014).

As mentioned in section 2.2.3, visual search strategy can be dichotomously classified into random or non-random. Once the search is non-random the description consistent is usually applied. Better target identification performance should be exhibited by inspectors using systematic as opposed to random visual search strategy (Nalangula et al 2006). It is also stated that the only way an exhaustive scan can be carried out, is by using a consistent search behaviour which ensures all areas are included in the inspection process (Melloy et al (2000). It is further stated that that a systematic search strategy is associated with inspectors exhibiting good visual search performance, and that any process that can improve the spatial uncertainty of visual targets will increase VIP (See, 2012; Eckstein, 2011).

For example, in a study of 206 professional baggage security searchers versus 93 non-professional (undergraduate) searchers, Biggs & Mitroff, (2013) found that search consistency predicted the accuracy of the professional searchers. They also found professionals were less affected by visual salience. With non-professionals searchers, accuracy was predicted by search speed as those searchers who took longer at the task were found to have the better visual search accuracy. Extrapolating to "real word" search scenarios, Biggs & Mitroff, (2013) suggest that there are various factors that can reduce search accuracy including multiple targets, inconspicuous targets, and unknown number of targets. In these situations, consistency of search behaviour may become an especially important factor in VIP and work by reducing the cognitive burdens placed on searchers during target identification.

2.7.9 Automated Visual Inspection

The field of automated visual inspection is a burgeoning field that has been succinctly summarised by Jiang et al, (2000) into the advantages and disadvantages of human versus machine inspectors. These scholars state that machines are more reliable, but humans are more creative and able to identify underlying patterns. Jiang et al, (2000) concludes by stating that both human and machine systems have their place in visual inspection. It is reported that automated inspection can in many cases replace visual inspection at a fraction of the cost and with improved accuracy. However, this improvement only occurs in a limited number of situations (Gallwey, 1998b). Furthermore human inspection is still considered better than automated inspection due to better pattern recognition, decision making and flexibility of humans (Rao et al, 2006). These scholars also state that humans will continue to play a very important role in “hybrid” human & machine inspection systems. Similarly Eckstein, (2011) states that “visual search by humans who are fully engaged by the search task, are yet to be surpassed by computer derived methods”. In short, visual inspection conducted by humans whilst being challenged by automation, may not be replaced any time soon.

2.8 Summary of the Literature

A succinct encapsulation of the literature relevant to visual search can be presented by quoting See, (2013). She states that human visual inspectors are imperfect, and exhibit large differences in performance across a range of inspection tasks. Visual inspection can be further summarised by saying that it is an error prone task, that is not easy to do well. The many reasons for this flawed human performance are listed in Table 3 and a reading of this literature can be interpreted as saying; it is of little surprise that humans should be so flawed when conducting visual inspections. However, VIP can be improved with methods put forward in the literature involving training, feedback and practice.

CHAPTER 3 VISUAL INSPECTION LITERATURE

3.1 Introduction

This chapter will focus on those published studies that have empirical evidence regarding either “real world” visual searches or the effect of eye scanning patterns on visual inspection performance (VIP). The chapter will begin by explaining the overall thesis literature search process by which a total of 95 studies were selected for use. It will then show how 43 of these 95 studies, were selected for this particular chapter. This will be followed by an explanation as to how nine of these 43 studies were then selected for in-depth analysis. For presentation reasons, these nine studies are divided into six field based studies with empirical VIP evidence, and three studies relating to the effects of set eye scanning patterns on VIP. These nine studies are also summarised in table format. The chapter will finalise with descriptive data from the remaining 34 studies that further evidence VIP conduct from the wider EHS community.

3.1.1 Literature Search Methods Used

In the first instance, both general and specific search methods were used in order to review potential published studies for inclusion in this thesis as whole. The general, more informal methods included viewing the reference sections of studies using the Google Scholar search engine with the terms “visual search” and “visual inspection”, reviewing specific authors associated with visual search and inspection and finally reviewing those journals dealing with the discipline. The specific, or more formal search methods used involved the Scopus Search engine which generated the descriptive statistics for the Prisma diagram in Table 5 below as described by Moher et al, (2009).

Scopus is an Elsevier product designed to be used and a single database and presents itself as the largest database available for peer reviewed literature searches (Scopus 2017). Coverage consists of nearly 23,000 journal titles from 5,000 publishers This search engine was used due to its inclusivity and ease of use.

3.1.2 Literature Search Inclusion & Exclusion Criteria

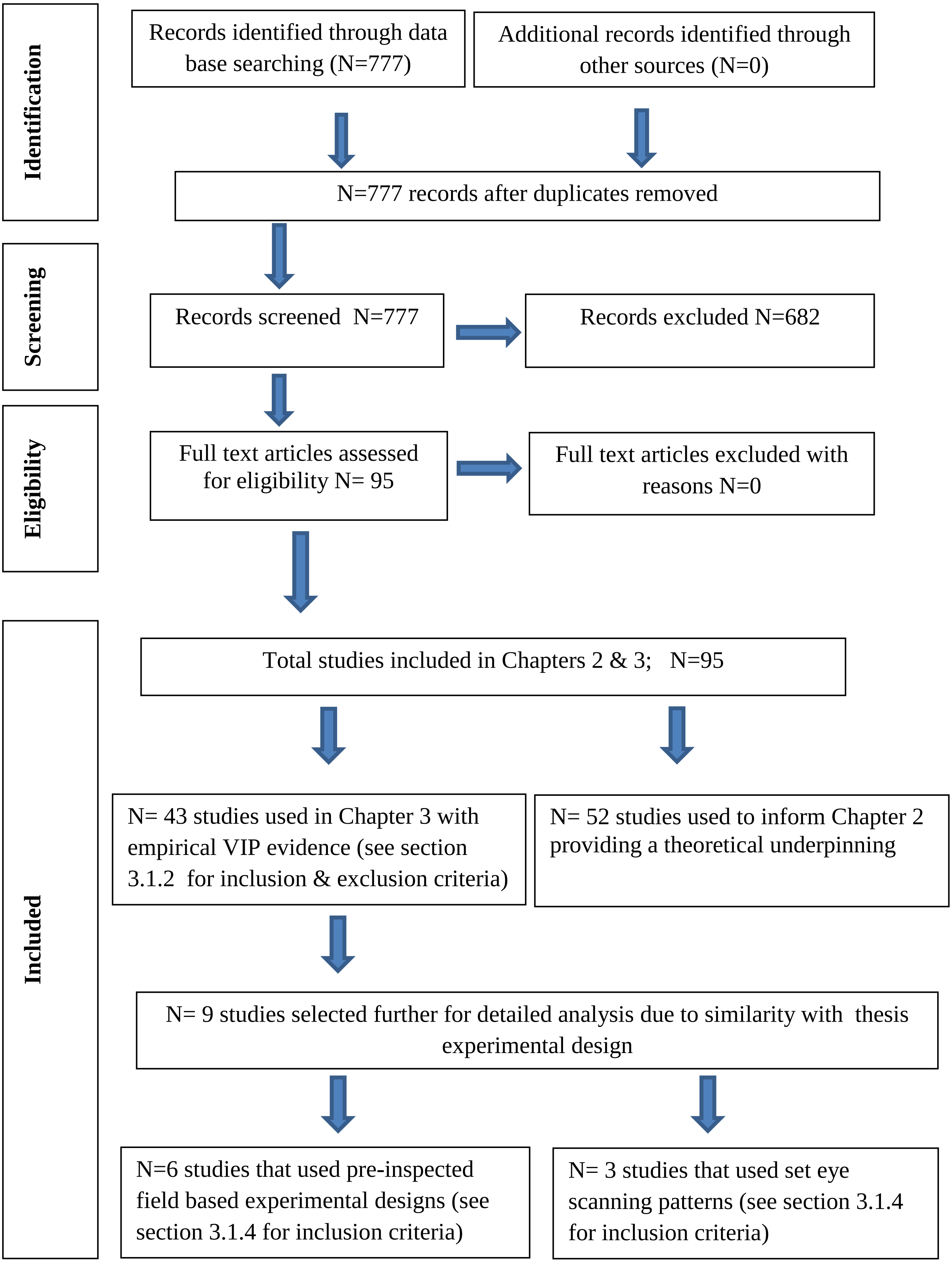
Regarding inclusion criteria for the entire thesis, a Scopus literature search for the terms “visual search” and “visual inspection” was entered. The total number of studies returned using the Scopus search engine for these two phrases was 777. In terms of exclusion, a total of 682 studies were excluded as they did not directly inform this thesis. A listing of the exclusion criteria used on these 682 articles is presented below;

- Visual search was part of a method, accepted as accurate and used to generate inquiry data. For example, virtual reality, using visual search to observe the number of cyclists not wearing helmets, or the accuracy of colour changes for chemical tests.
- The inquiry revolved around the utility of visual search applications, but without any evidence of VIP, for example drone mediated inspection of tall buildings.
- The research involved single or low number targets under laboratory conditions, to inform attentional inquiry or visual neuroscience.
- Visual search was modelled mathematically for machine learning, artificial intelligence or computer aided inspection applications.

By far the largest types of study excluded from these 777 publications were those that focussed on visual search attentional inquiry or visual neuroscience, using a single target laboratory based research methodology. The vast majority of visual search inquiry involves this type of methodology (Eckstein, 2011; Nakayima & Martini, 2011) hence the large number of studies excluded. In contrast, the research question asks; "why are visual inspections often flawed and how can these flaws be overcome". This question requires inquiry with a focus on the multiple observation of multiple hazards within a "real world" setting by EHS professionals conducting visual inspections. Therefore, this type of visual search attention or visual neuroscience inquiry was of limited value to the thesis aim, hypothesis and question. However those cognitive visual psycho-physics review articles, which speculate on applying "real world" visual search inquiry, but still based on attentional inquiry or visual neuroscience, are included. Three such review articles have been used in this thesis, Eckstein, (2011); Nakayima & Martini, (2011); See, (2012).

Having excluded 682 studies, there were 95 published studies left that have been reviewed and used to inform this thesis. Of these 95 studies, 52 studies have informed Chapter 2 (Theoretical Underpinnings). Nine of the remaining 42 studies, that specifically evidence EHS VIP and the effects of eye scan patterns have been selected for detailed analysis in this Chapter. The 34 related studies left, that did not so closely resemble the experimental design used in this thesis are presented in Appendix 8. The Prisma diagram in Table 5 below, summarises the literature search process used.

Table 5 Literature Search Prisma Diagram



3.1.3 Trends in Visual Search & Inspection Inquiry

The historical development of visual inspection research has generally developed from two distinct disciplines. The first discipline is laboratory based visual psycho-physics which has largely involved visual search scientists and psychologists. The second discipline is industrial engineering (usually applied to a quality control function), which has largely involved engineers (Eckstein, 2011; Nakayima & Martini, 2011; See, 2012).

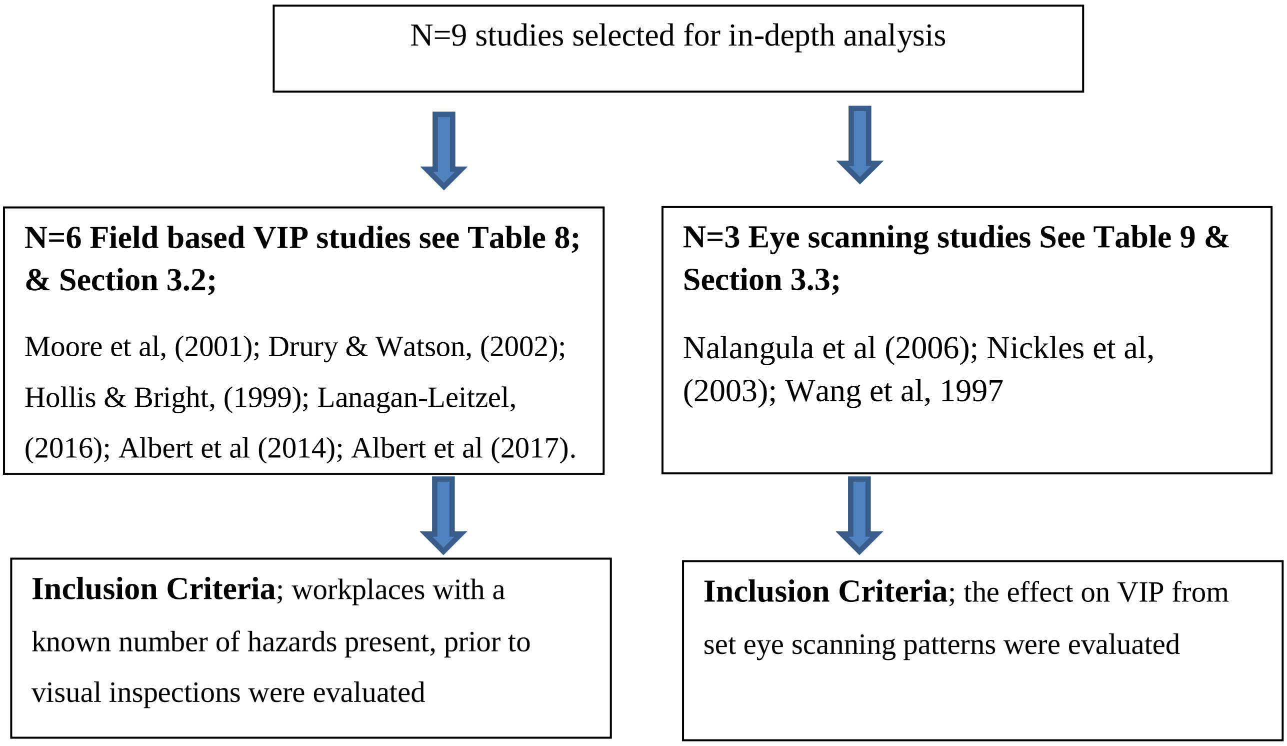
Of the two disciplines, visual search psycho-physics has received far more academic attention with the literature in this field being described as voluminous by Wolfe, et al (2005). To illustrate this point, the visual search scholars Nakayama & Martini, (2011) stated that the number of published studies with a visual search psycho-physics background was in the order of 2,200 articles. In contrast See, (2012) puts the number of publications from the visual inspection literature from the industrial engineering discipline during the last six decades at 212. This is circa one tenth the output.

The evolution of the specific discipline of visual inspection from an industrial engineering discipline, did not begin until the 1950's, even though there has been a long tradition of visual inspection research going back to the 1930's (Nickles et al, 2003; See, 2013).

3.1.4 Inclusion Criteria of Nine Studies Presented for Detailed Analysis

Having identified 43 studies that empirically evidence the VIP of EHS professionals, nine studies were selected for further in depth presentation and critical analysis. The selection criteria for this step will now be explained. These nine studies have one of two inclusion criteria. The first criterion is similarity to the thesis experimental design. All these published studies involved EHS professionals who were tested under ecologically valid, "real world" conditions and in workplaces with a known number of hazards. This pre-determination of the number of hazards in the area under analysis, allowed for precision measurement of VIP. Six such studies meet this inclusion criterion. The second criterion is the effect of set eye scanning patterns on VIP. Three such published studies met this criterion. As can be envisaged, these nine "core" studies naturally further divide into; six studies which relate to field based VIP inquiry and three studies that relate to the effects of eye scanning patterns on VIP. A summary listing of these studies listing their categories and inclusion criteria is given in Figure 2 below;

Figure 1 Categorisation & Inclusion Criteria for the Nine Studies


N=9 studies selected for in-depth analysis

N=6 Field based VIP studies see Table 8; & Section 3.2;

Moore et al, (2001); Drury & Watson, (2002); Hollis & Bright, (1999); Lanagan-Leitzel, (2016); Albert et al (2014); Albert et al (2017).

N=3 Eye scanning studies See Table 9 & Section 3.3;

Nalangula et al (2006); Nickles et al, (2003); Wang et al, 1997

Inclusion Criteria; workplaces with a known number of hazards present, prior to visual inspections were evaluated

Inclusion Criteria; the effect on VIP from set eye scanning patterns were evaluated

This subdivision has been implemented to allow a comparative critical analysis between these nine studies and this thesis, in a clear and concise manner. This analysis will present a summary listing of the two categories of core studies listed in Figure 2 above and as can be seen in Table 8 & 9. A further notable finding from this literature review to consider is that only six published studies have researched field based VIP by EHS professionals, and three for the effects of using consistent eye scanning patterns on visual inspection performance. This illustrates how limited the current level of "real world" visual inspection research actually is. Therefore, there is a lack of underpinning evidence for the thesis aim and research question in the wider visual search literature which Chapter 5 (Results) will attempt to rectify.

3.2 The Six Field Based VI P Studies Included

This section will now present those six published studies that most closely resemble the experimental design used in this thesis. It will begin by detailing four published studies whereby EHS professionals whose day to day job involved observing workplace hazards were quantitatively evaluated. The first four presented are; the frequency of construction defects hazards observed by bridge inspectors (Moore et al, 2001) the probability of aviation engineers observing airframe defects (Drury & Watson, 2002) and the VIP of ten building surveyors who all visually inspected the same dwelling house (Hollis & Bright, 1999). There is a further study included here; Lanagan-Leitzel, (2016), which involved allowing mannequins to float to the bottom of swimming pools and timing the reaction of attending lifeguards. This study is from the grey literature but it is included here due its relevance and perceived high standard of scholasticism by the lead investigator. It is cited in a peer reviewed journal, which also increases its relevance for inclusion. Two further studies are then presented that quantify how many hazards were observed by construction site workers (Albert et al, 2014; Albert et al, 2017). These two studies evidence the VIP of construction site workers who had visual inspection duties but were not professional searchers. This is a distinction to bear in mind. These construction workers would still be expected to be observe and act upon construction site hazards. However, this distinction moves them slightly away from the research question that involves EHS professionals whose sole function during inspections is looking for workplace hazards. The six studies presented here, all empirically quantify “real world” VIP of EHS professionals. These six studies, as summarised in Table 8, are now explained in further detail below.

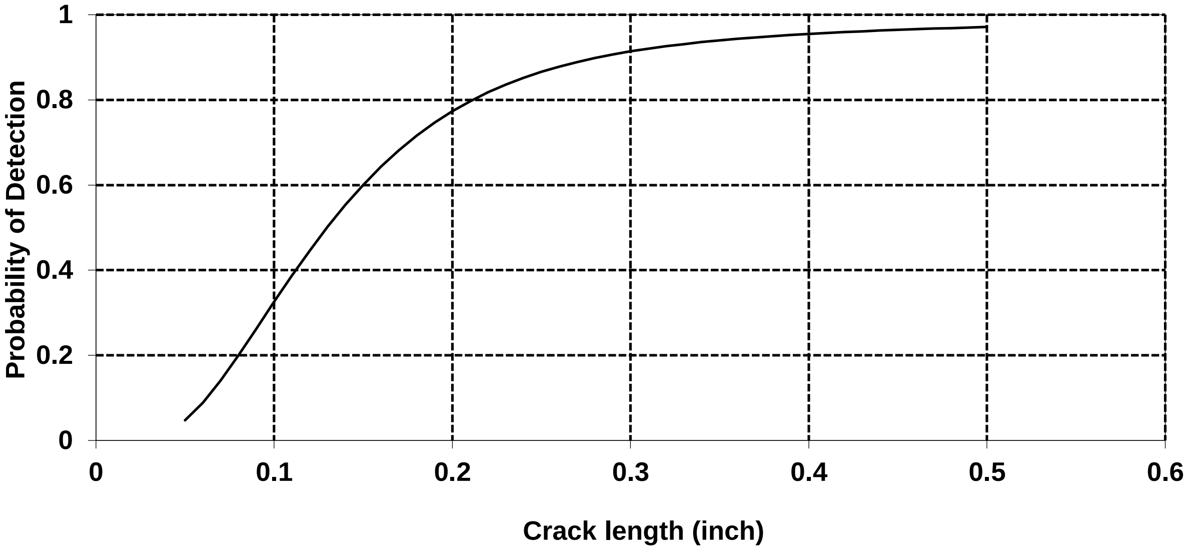
3.2.1 The Bridge Visual Inspection Study

The Moore et al, (2001) research team evaluated the VIP of US licenced, transport and highway bridge inspectors who were tasked with conducting bridge safety reports on pre-inspected bridges. In this study, 49 engineers were recruited and 37 bridges were used. Each engineer visually inspected ten bridges and completed six routine and two in depth visual inspections. These visual inspections were evaluated by their peers whilst on site. Each of these bridges had been pre-inspected by the Moore et al, (2001) research team beforehand. This produced a listing of the engineering defects and faults on the thirty seven bridges that should have been observed. The bridge inspectors were tasked to inspect their assigned bridges using their normal custom and practice. The results demonstrated a great deal of variability in terms of VIP. Less than 10% of bridge inspectors indicated the presence of skew, whereby a bridge component bends or moves out of its intended alignment. Less than 25% indicated defective support conditions and less than 50% indicated fracture critical conditions. The conclusion here is evidence, once again, for observable hazards not being seen during visual inspections.

3.2.2 The Aircraft Maintenance Visual Inspection Study

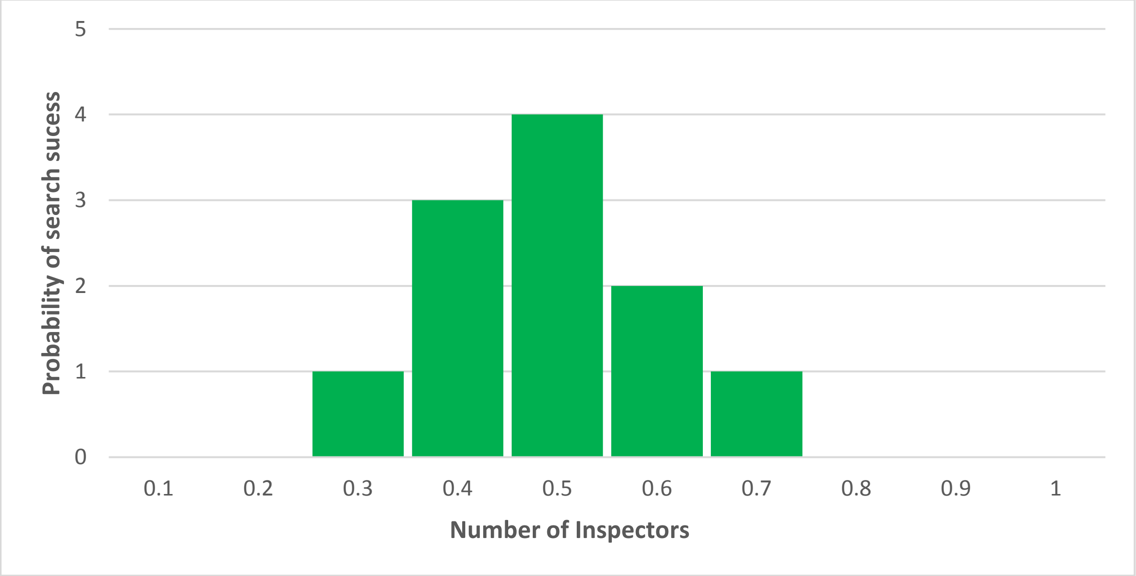
Drury & Watson (2002) present a cited study from Spencer, et al, (1996). The data resulted in a probability of detection curve for cracks observed on one Boeing 737 aircraft. The participants were 12 experienced aircraft maintenance technicians who conducted visual defect inspections. The aircraft in question was pre-inspected by the research team for defects prior to these visual inspections. The resulting probability of detection by the inspectors varied between .05-.95 depending on crack size. This data is presented as a receiver operating characteristic (ROC) probability curve whereby the probability of detection was graphed against airframe crack sizes. ROC curves are a standard engineering method of displaying performance characteristics. A perfect VIP whereby all defects are observed would produce a straight line ROC "curve". A negative deflection from this straight line quantifies the error rate and presents a graphical display of how much visual inspection error was demonstrated. The ROC is shown in Figure 2 below.

Figure 2 Probability of Detection Curve



The same cited study also presents a probability density function shown in Figure 3 below.

Figure 3 Probability Density Function for Visual Inspector Search Success



The ability to interrogate the data from both the bridge and aviation study is hampered. This is due to the way the data is presented and in particular, with the overall number of hazards given in probability terms and not as a whole number. This is (probably) deliberate and understandable, given the sensitivity of any published data relating to the safety record of both the US bridge maintenance programmes and the US aviation industry.

3.2.3 The Building Surveyor Visual Inspection Study

In this study, a two story dwelling house in London was used. This building was pre-surveyed by the study team to reveal six specific housing defects. All these defects would be expected to be observed by the study participants being ten surveyors. These participants were paid by the Hollis and Bright, (1999) research team to produce house condition surveys for prospective purchasers. These surveyors were unaware that they were participants in this study. The research team presented themselves as prospective purchasers for the house to each of the ten surveyors who participated in this study. The defects that should have been observed in the house were; an unsupported chimney in the roof void, subsidence in the rear extension, defective rear elevation window frames, water entry due to defective gutters, water entry underneath window sills and defective rendering. Each of these ten surveyors was representative of the industry, conducting such surveys on a day to day basis. All were professionally qualified to conduct dwelling house conditions surveys.

Of the ten surveyors in this study, only one identified all six defects. One surveyor identified five defects, three surveyors identified four defects and one surveyor did not identify any defects. The remaining five surveyors only identified two defects. These scholars speculate that it was time spent surveying that predicted VIP with the "worst" three surveyors spending a mean 1 hour 7 minutes on site. The remaining seven surveyors spent a mean 2 hours and 12 minutes on site.

3.2.4 The Lifeguard Visual Inspection Study

A final study is included here, Lanagan-Leitzel, (2016) that involved swimming pool lifeguards. As it is from the "grey" literature, the reliability and validity of this study is debatable. Although it does not meet the selection criteria of a peer reviewed journal study, it is nevertheless included here as it is a plausible account of VIP inquiry. It is from a trade journal, albeit with credible sources. Therefore readers are invited to judge for themselves, the credibility of the Lanagan-Leitzel, (2016) cited study, by Brener & Oostman, (2002). This study reported that an international water safety consultancy company tested the VIP of lifeguards at 90 swimming pools in the USA. In this study 500 tests were conducted whereby mannequins were placed into pools and allowed to sink. The subsequent actions of the attending lifeguards were then noted for analysis. In 14% of these tests, the lifeguards failed to notice the mannequins at the bottom of the pool within the required industry standard of three minutes. Furthermore, in fewer than 10% of tests did lifeguards act in accordance with industry standards that stipulate; drowning incidents should be observed within 10 seconds and a rescue commenced within 20 seconds of the incident beginning.

3.2.5 VIP Training for Construction Workers Using a Mnemonic

Two further studies will now be presented. The Albert et al, (2014) and Albert et al, (2017) studies evaluated the VIP of construction site workers before and after a training intervention.

In the Albert, et al (2014) study, the authors recruited 78 construction workers and evaluated how many workplace hazards they normally observed, (or in the authors words; "recognise and communicate") whilst working on construction sites. Having evaluated their normal custom and practice, the study team used a training intervention aimed at improving their VIP. Six work crews and two large construction sites were used with three crews per site. A total of 96, 4 hour work periods were evaluated. The sites were valued at USD 13 Million and 550 Million respectively, demonstrating that these were large construction projects.

The training intervention consisted of construction workers being trained in how to better observe hazards on site by using a 10 term mnemonic. Hazards were categorised according to the type of energy they represented. The workers were trained to remember possible site hazards by the use of this mnemonic which was intended to remind participants of hazards whilst on site. The terms used for the mnemonic intervention were; gravity, motion, mechanical, pressure, radiation, biological, chemical, sound, electrical, and temperature. For example participants were expected to improve their observation of possible falling overhead objects by remembering the term gravity. Participants were expected to improve their observation of site transport machinery by remembering the term motion, and so on. This intervention was therefore aimed at improving VIP by increasing recall of their attentional hazard set which in turn, was intended to improve the observation of hazards when back on site. In effect, the mnemonic was used to create a mental checklist for construction workers in order that they better recall, and thereby better observe, relevant site hazards.

The actual number of hazards present on the sites was arrived at by the use of 14 industry experts and senior site managers who used two methods. Firstly, they brainstormed the type of hazards to be expected whilst the crews were working. Secondly, they conducted inspections themselves whilst the crews were working. Subsequently, the number of hazards the construction site workers recognised and communicated, was used to calculate VIP of the workers. The term recognised and communicated was not explained in their paper, but is taken as meaning hazards were observed before being written down or verbalised. Albert et al, (2014) reported that prior to the intervention, participating construction workers observed a mean 38% of construction hazards. After the 10 term mnemonic training intervention, participating construction workers were again taken to the same construction site, and asked to observe the hazards present. The level of VIP subsequently increased by approximately 12%, to a mean of 49.75%.

Albert et al (2017) extended this methodology and collected data from 100 subsequent construction workers, in 18 different crews and on 6 different US construction sites. Three further interventions were introduced, each incorporating the same 10 term mnemonic training method as described above from their 2014 study. The difference in this 2017 study was that the mnemonics based training was delivered using three separate pedagogies. The first intervention used virtual reality to depict site hazards and was called the "system for augmented virtuality environment safety" (SAVES). The second intervention used brainstorming during site meetings to identify hazards and was called the "safety meeting quality measurement maturity model" (SMQM). The third intervention used work site displayed graphical representations of construction hazards known as the "hazard identification and transmission technique" (HITS). VIP prior to intervention by was 32.30%. Using the mnemonics approach but with the three different educational delivery methods, increased the overall rate of VIP quite spectacularly, and in each case to over 70% as Table 7 below demonstrates.

Table 6 VIP from 3 Training Interventions

	VIP at the Pre-Intervention Stage	Resulting VIP Using the SAVES Intervention	Resulting VIP Using the SMQM Intervention	Resulting VIP Using the HITS Intervention
Mean percentage hazards observed	32.30	72.60	70.20	83.00

It can be argued that a direct implication of the Albert et al, (2017) study is that the HITS intervention, whereby pictures of hazards are simply shown to construction workers, is far more successful than SVS and should therefore be used. However it should be borne in mind that all the interventions in the Albert et al, (2017) study relate to what hazards to look for. SVS in sharp contrast, is researching how to look for hazards. As will be seen in the results chapter and in particular section 6.4.6, how to look for hazards fundamentally affects visual inspection performance and is central to the inquiry in this thesis. Furthermore sections 6.5.6 and 6.5.7 discuss why SVS did not achieve the findings reported in the Albert et al, (2017) study.

3.2.6 A Comparison and Critique of these Six Field Based VIP Studies

The similarities and differences between these six studies and this thesis will now be described (see also section 3.1.4 on inclusion criteria).

There are similarities between these six published studies and this thesis. Firstly, all the participants can be described as EHS professionals who routinely conduct visual inspections. Secondly, the VIP of all the participants was quantitatively evaluated under ecologically valid conditions. The obvious difference is that other than aviation engineers, the remaining categories of professionals from these published studies were from different discipline backgrounds. In addition the thesis methodology involved a randomised controlled trial whereas all these studies were correlational or interventional in nature (see Table 8).

Regarding the bridge, aircraft, and building studies, the sample size of participants was too small for statistical power purposes. The largest of these studies, the bridge inspection study by Moore, et al (2001) used 49 inspectors conducting ten inspection tasks each. This is still below the 80 cases recommended by Field, et al (2013) for correlational research to achieve sample power. The two remaining studies Drury & Watson, (2002); Hollis & Bright, (1999) used 12 and 10 inspectors respectively. Whilst small sample size does not detract from the quantitative results of VIP demonstrated, they cannot by convention be generalised (Field, 2013).

The main drawback of the Lanagan-Leitzel, (2016) cited study by Brener & Oostman, (2002) is that it is from the “grey” literature. This lack of academic peer review oversight cannot be ignored, but neither can the results. This cited study is compromised but nevertheless represents further, but weaker quantitative evidence to show just how difficult it is for EHS professionals to observe all relevant workplace hazards during inspections.

The next two studies were both conducted by the same research team; (Albert et al, 2014; and Albert et al, 2017). A material difference between these two construction based VIP studies and the four studies above; (Moore et al, 2001; Drury & Watson, 2002; Hollis & Bright, 1999; Lanagan-Leitzel, 2016) is that the prior knowledge and educational status of construction participants was not reported. Whereas the EHS professionals in the remaining four studies, would have all been qualified to visually inspect.

Sample size was higher than the 80 cases as recommended by Field, 2013 to achieve statistical power. However, both studies lacked any data on the actual number of hazards that were on site pre or post intervention. In the Albert et al, (2017) study in particular very high final VIP results are claimed with a minimum of 70% of all site hazards observed due to their mnemonic training interventions (see Table 7). However, if the frequency of hazards on site were in single figures to begin with (for example sound was classified a hazard) then these high levels of improvements could be explained by chance or by a small increase in relative frequency. Unfortunately, a statistical treatment to assess significance or effect sizes of the improvements was not given for individual hazards in this study.

A further critique of the Albert et al, (2013) and the Albert et al, (2017) studies is the way in which the actual number of construction site hazards was arrived at. In both studies the number of hazards on site was ascertained using a combination of managers visually inspecting the site and scenario analysis. Regarding the managers visual inspection methods, the same causes of visual search error as summarised in Table 3 could have resulted in a lesser number of hazards that actually existed, thereby possibly inflating their success rate. This point can also be raised with the remaining four studies (Moore et al, (2001); Drury & Watson, (2002); Hollis & Bright, (1999); Lanagan-Leitzel, 2016). However, the key difference here is that these four studies all had researchers who were professionally qualified to visually inspect.

Furthermore, the two Albert et al, studies did not use a control group. As these studies represent a pre-post intervention experimental design, any motivational effects cannot be ruled out. It is possible therefore, that some of the high levels of VIP improvements could be reflective of participants trying harder to please the research team. This is a possibility whenever a pre-post experimental design is used without a control group (Aronson, 2012).

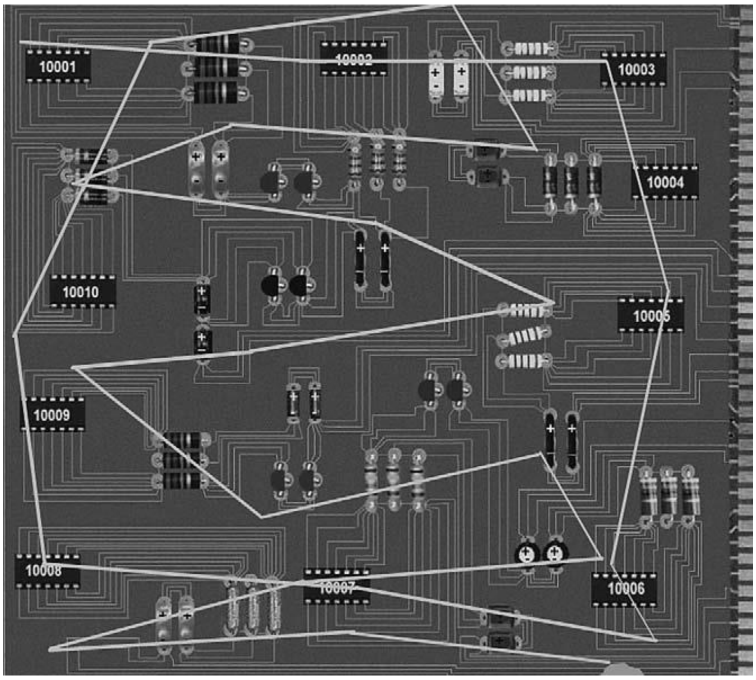
3.3 The Three Eye Scanning Studies Included

There are three published studies that researched the effect of eye scanning patterns on VIP. These studies in order of presentation are; Nalangula et al (2006); Nickles et al, (2003) and Wang et al, 1997. These three studies are presented here primarily because the intervention used in this thesis also included a set eye scanning pattern to observe workplace hazards. These eye scanning studies will now be presented in detail and are summarised in Table 9.

3.3.1 The Nalangula et al, (2006) Study

This study involved 32 novice students who were tasked with looking for defects on a PC displayed printed circuit board (PCB). These students were randomly assigned to four treatment groups; static, dynamic, hybrid and control. Other than the control group, they were all trained in the use of a set eye scanning pattern. The eye scanning pattern used was to start observing in the top left corner of the PCB display, and then scan all the edges in a clockwise pattern before a reverse snakes and ladders scan path down the remaining PCB area. This pattern can be seen in Figure 5 below.

Figure 4 Eye Scanning Pattern Used by Nalangula et al (2006).

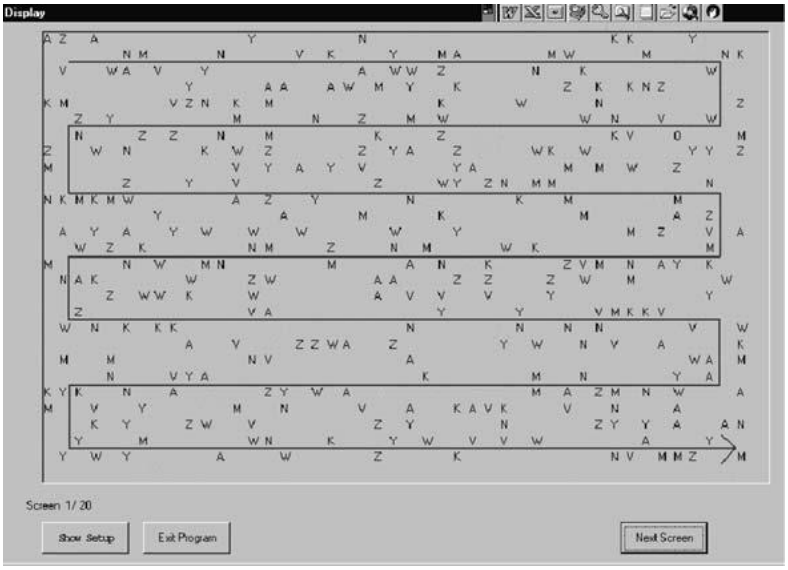


The static group receive a screen shot picture showing the eye scanning pattern used in Figure 4 above. The dynamic group were shown the same screen shot on the PC, but with a yellow cursor directing the eye over the same path. The hybrid group received the PC displays screen shot and cursor, but with the addition of a static trace of the area followed by the cursor. The control group were just directed just to look for defects on the PC display, with no training in using an eye scanning pattern. The main findings reported by these scholars were the positive effects of using a set eye scanning pattern when searching, and that search strategy can be taught.

3.3.2 The Nickles et al, (2003) Study

This study involved 24 college students as participants, who were tasked with finding letters on a PC screen using a set eye scanning pattern. What was being evaluated in this study was whether the method of instruction for a set eye scanning pattern, would have any effect on VIP. Each participant was tasked to use a "reverse snakes and ladders" observational path as seen in Figure 6 below.

Figure 5 Reverse Snakes & Ladders Eye Scanning Pattern by Nickles et al, (2002)

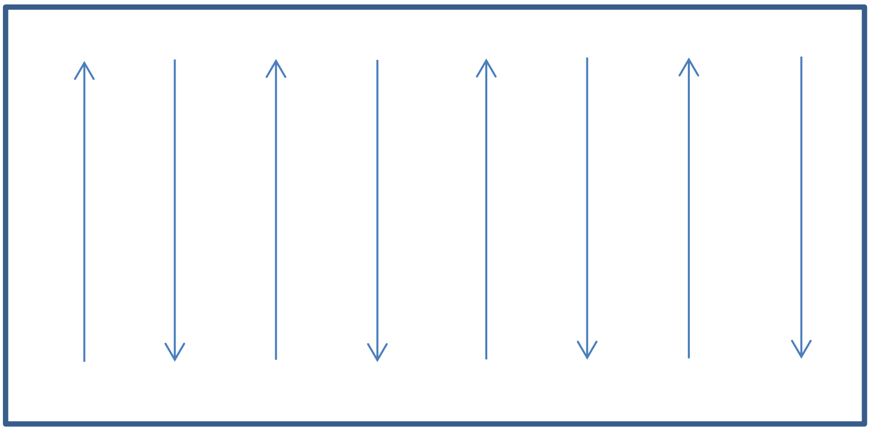


These 24 participants were divided into three groups. The first group were trained using verbal instruction to use a reverse "snakes and ladders" eye scanning pattern as seen in Figure 6 above. The second group were trained using with verbal instruction and a diagram showing this pattern. The third group were trained using the same diagram, but with accentuated graphics on the PC display that traced the required pattern. It was demonstrated in this study, that how the set eye scanning pattern was taught made little difference. It was the use of the set eye scanning pattern itself, that improved VIP when compared to random visual search behaviour. This study further demonstrated that mean inspection time decreased by 49.9%, search time decreased by 59.6%, mean stopping time decreased by 46.1% and the mean number of fixations decreased by 46.3%. Furthermore, there was no significant difference between the type of intervention used. These scholars concluded that the simplest form of training for SVS was verbal, and therefore should be adopted.

3.3.3 The Wang et al, (1997) Study

This study involved 30 students who were tasked with finding soldered joint defects on PCB's that were displayed on a PC screen. Ten students were divided into groups that used either systematic, natural or random search patterns. The ten systematic group participants were trained in a consistent eye scanning pattern that followed an alternating up and down left to right observational path as seen in Figure 6 below.

Figure 6 Up & Down & Left to Right Eye Scanning Pattern after Wang et al, (1997)



The ten participants in the natural search group received no search strategy training but did receive feedback on performance after each screen display. The remaining ten participants in the random search group were instructed not to use a fixed or regular search strategy. The results in percentage defect detection rates were; systematic search, 96%, natural search, 83%, and random search, 81%. These results were significant ($p=\leq.05$). The authors concluded that a set eye scanning pattern is beneficial for VIP and can be trained.

3.3.4 A Comparison & Critique of these Three Set Eye Scanning Studies

There were material differences in the methodology used in these three published studies and this thesis. The targets required to be observed in these three studies were all displayed via PC screens and can be described as lab experiments. The set eye scanning pattern in the two thesis studies differs, and was applied under far more ecologically valid settings involving EHS professionals visually inspecting workplaces. The eye scanning pattern used by participants in the two thesis studies was the "reading a book" pattern (see Figure 1). In contrast, Nickles et al, (2003) used a reverse "snakes and ladders" eye scanning pattern.

Wang et al, 1997 used an up and down, left to right eye scanning pattern. Nalangula et al 2006 used an outer borders scan first, before a reverse “snakes and ladders” eye scanning pattern for the remaining area under analysis. Finally these three studies required participants to look for a narrower range of targets being either; defects in PCB’s or certain letters. In contrast, participants in the two thesis studies were required to look for a much wider range of hazards.

These three published studies were also relatively small scale with 32 participants per study at the most. In contrast, the total number of thesis participants in the kitchen and aviation studies was 211 and 26 respectively. Therefore, by convention these three published studies cannot be generalised (Field, 2013). The headline however, from these three studies is the beneficial effects on VIP from using a set eye scanning pattern when compared to “normal” observational behaviour. These three studies also represent the only research so far published that specifically evaluates the effect of using a set eye scanning pattern on VIP. These three studies are therefore important for this thesis, as they represent the only examples of peer reviewed quantitative data to show how VIP can be improved when a set eye scanning pattern is used. They also show that more research is required to evaluate the effect of set eye scanning patterns, when applied to “real world” visual inspections of workplaces.

Table 7 Summary Listing of Field Based VIP Studies

Author	Title	Type of Study	Method	Main Findings
Moore, et al, (2001)	Reliability of Visual Inspection for Highway Bridges	Correlational	N= 49 inspectors conducted ten inspection tasks on seven bridges drawn from a list of 37 bridges. All the bridges had been pre-inspected by the research team.	• Significant variability reported. • Using a 1-5 scale bridge condition rating by inspectors, 68%*** of ratings varied within 1% point of the average and 95% of these ratings varied by two points of the average. • At least 56% of condition ratings were found to be incorrect
Drury & Watson, (2002)*	Good Practices in Visual Inspection	Correlational	N=12 aircraft maintenance technicians visually inspected a Boeing 737 that had been pre-inspected by the research team	• Probability of search success range was .3 to .7 • The probability of search success mode was .5
Hollis & Bright, (1999)	Surveying the Surveyors	Correlational	N=10 surveyors each inspected the same dwelling house with six known housing defects, pre-inspected by the research team	• One surveyor observed all six defects • One surveyor did not observe any defects • Five surveyors observed two defects only • Time on site is given as the reason for the results with some surveyors using twice the inspection time

Lanagan-Leitzel, (2016)**	Great Expectations: Perceptual Challenges of Visual Surveillance in Lifeguarding	Intervention	N=500 tests were conducted in 90 swimming pools whereby mannequins were allowed to sink and lifeguard behaviour noted.	• In 14%*** of these tests, lifeguards failed to notice the mannequins at the bottom of the pool within the industry set standard of three minutes, • lifeguards met industry set standards in less than 10% tests being; drowning incidents should be observed within 10 seconds and a rescue commenced within 20 seconds.
Albert et al, (2014)	Enhancing Construction Hazard Recognition and Communication with Energy-Based Cognitive Mnemonics Safety Meeting Maturity Model: Multiple Baseline Study	Intervention	N=78 construction workers were asked to identify workplace hazards. A training intervention based on a ten term hazard mnemonic was used to improve on site hazard identification. The actual number of construction hazards present during work periods was assessed by site safety professionals.	• Baseline hazard observation was level found to be 38%*** of the total number of hazard present. • The mnemonic training intervention achieved a VIP improvement to 49.78%

Albert et al, (2017)	Empirical measurement and improvement of hazard recognition skill	Intervention	N= 100 construction workers were asked to identify workplace hazards. The same 10 term mnemonic training intervention was used by Albert et al, (2014) but delivered by three different methods; virtual reality, brainstorming and hazard graphics. The actual number of construction hazards present during work periods was assessed by site safety.	• Baseline hazard observation was found to be 32.30%. • Each mnemonic based training intervention achieved a subsequent improvement to over 70% of hazards present on site observed. (see Table 7 for full details)

***cited from Spencer, Schurman and Drury (1996).**

**** classified as “grey” literature and cited from Brener & Oostman, (2002).**

***** As reported directly by the study.**

Table 8 Summary Listing of Eye Scanning Pattern Studies

Nalangula et al, 2006	Evaluation of the effect of feedforward training displays of search strategy on visual search performance	Randomised Controlled Trial	N=32 novice students were randomly assigned to four treatment groups; control, static, dynamic and hybrid. Excepting the control group they were all trained in the use of a set eye scanning pattern, (see Figure 4). They were tasked with looking for defects on a PC displayed printed circuit board. The static group receive a picture showing the eye scanning pattern. The dynamic were shown the same but with a yellow cursor directing the eye over the same path. The hybrid group received all above and a static trace with a cursor. The control group looked for defects with no training.	• The mean percentage observation rate of each group was as follows; • Control; 61.58% (SD=8.55) • Static; 67.46% (SD=11.48) • Dynamic 72.42% (SD=6.02) • Hybrid; 78.12% (SD=8.29) • The results were significant indicating the positive effects of using a set eye scanning pattern when searching and compared to a control group. • The results demonstrated that search strategy can be taught • The hybrid method was not significantly different to the dynamic method
Nickles et al, (2003)	A comparison of three levels of	Intervention	N=24 subjects were tested with finding letters on a PC screen. They created	As a result of the interventions; • Mean inspection time decreased by 49.90%,

	training designed to promote systematic search behaviour in visual inspection.		three groups each of which received training in a set eye scanning pattern; following a "reverse snakes and ladders" observational path (see Figure 5). The first group were trained using verbal instruction. The second group with verbal instruction and a diagram and the third group had the same diagram but with accentuated graphics tracing the eye scanning path	• Mean search time decreased by 59.60% • Mean stopping time decreased by 46.10% • Mean number of fixations decreased by 46.30% • There was no significant difference between the type of intervention used. These scholars conclude that the simplest form of training, being verbal, should be adopted.
Wang et al, (1997)	Training for Strategy in Visual Search.	Intervention;	N=30 students were allocated a search task for finding soldered joint defects on printed circuit boards displayed on a PC screen. 20 students were trained in a consistent eye scanning pattern (see Figure 7) beforehand as well as the defect type. They were the allocated into either systematic, natural or random search pattern groups. In systematic search students received	The results in percentage defect detection were • Systematic search; 96% • Natural search; 83% • Random search; 81% • The authors conclude that a set search strategy can be trained. In this study, the strategy used was alternating up and down starting from the left and scanning to the right.

			training in a systematic search being alternating up and down left to right. Natural search students received no search strategy but did received feedback on performance after each screen display. In random search participants were instructed not to use a fixed or regular search strategy	

3.4 Summary of Visual Inspection Performance Studies

When considering VIP from the studies listed above, then a number of key findings emerge. Firstly, field based evaluations of VIP reveal their error prone nature. Secondly, regarding published field based EHS professionals generally low rates of VIP are reported being in the thirties range percentage wise. This VIP figure rises as the nature of the visual inspection task becomes more specialised and thereby involves less hazards. Thirdly the generally low level of visual inspection performance reported in the visual inspection literature has not drawn any major critical comment from the environmental health and safety community. This is a point made by Gallwey, 1998(b). In addition Woodcock, (2014) reports that the Environmental Health and Safety community simple assumes visual inspections are conducted and do not question their accuracy. However, concern has been raised by scholars working in specialised areas such as aviation maintenance, medical image diagnostics, industrial quality control and the transport security field (see for example Drury & Watson, (2002); Wolfe et al, (2005); Gallwey, (1998b); Biggs & Mitroff, (2013). Finally all three published studies that evaluated the use of set eye scanning patterns reported improvements in VIP.

In summary, visual inspection error is a widespread phenomenon occurring in whatever field VIP search inquiry is conducted. Examples of serious visual inspection failures can be found in various domains including construction engineering, industrial inspection, security screening and medical images diagnostics. Given these results, there is evidence to show that flawed visual inspections are likely to occur, whenever visual inspections are required.

CHAPTER 4 METHODOLOGY

4.01 Introduction to the Methodology

This chapter will present the methodology used in this study. It will be broken down into an evolution of the problem, the epistemology followed, the method used, and finally a statistical analysis plan. The first section will begin by laying out how the lead investigator came to consider the research hypothesis, and thereby the epistemological philosophy underlying the methodology. It will then detail the arguments as to why a randomised controlled trial methodology with a high degree of pragmatism and ecological validity was considered the best approach for the research hypothesis. The next section will detail the recruitment of participants, the SVS method itself and the random allocation procedures used. The final section, the statistical analysis plan, will detail the mathematical treatment of the data generated and detail the descriptive and inferential tests used for subsequent statistical analysis.

4.02 Evolution of the Epistemological Philosophy & Research Question

Since he began his career over thirty years ago in London as an Environmental Health Officer, how to conduct visual inspections has long been a research interest of the lead investigator. His regulatory role was entirely predicated on visual inspection to inform appropriate enforcement responses. Typical legal enforcement procedures involved prosecutions of food service organisations for unsafe or unhygienic conditions. During such proceedings, it was the strength of his visual inspection evidence that supported his decisions as to whether or not workplaces were safe or required a regulatory response. A sobering effect of court appearances was the realisation that defence barristers were adept at finding limitations of visual inspection conduct. And rightly so. If the visual inspection evidence was not sufficient to show a workplace was unsafe or unhygienic, then the defendant expected his natural justice right of acquittal. These legal encounters reinforced a deeply held view that visual inspections need to be conducted in a competent and professional manner. In this way, sufficient evidence regarding safety would ensure any decision made was the appropriate regulatory response. However, what constituted a competent visual inspection was open to debate as there was little professional conduct guidance available.

Within the lead investigator's local authority, all Environmental Health Officers were trained in visual inspection conduct based on the Royal Institute of Chartered Surveyors method (see section 2.7.4). Even though the lead investigator did not find this method was required in other organisations he worked for, he nevertheless used the method for all future visual inspections.

Some twenty years ago, the lead investigator, took up a lecturing position in the Dublin Institute of Technology teaching and researching the discipline of EHS. It was a natural progression for the lead investigator to teach this method of visual inspection conduct to his MSc students. However, he soon found that very few of his students were aware of this type of visual inspection conduct. Furthermore, many students did not seem to be aware of other EHS professionals who conducted inspections in a similar manner. With this thought in mind, the problem began to crystallise into a research project. Here, the lead investigator became determined to show the value or not, of using his systematic visual search (SVS) behaviour during workplace inspections.

The first step in researching this issue began when the lead investigator attempted to evaluate how EHS professionals commonly conducted their visual inspections of workplaces. The resulting study (Hrymak al, 2015) recruited 40 experienced EHS professionals and demonstrated that there was no standard visual inspection behaviour amongst these professionals. Their visual inspection methods could best be described as idiosyncratic with large variations in what they looked for, how they looked, for how long they looked and in what order they looked. A finding from this qualitative study was that different EHS professionals would report different hazards for the same scenario under analysis. A further finding was that there is a scant body of knowledge for the adequate conduct of visual inspections. One further consequence of this study on the epistemological philosophy for this thesis, was the confidence exhibited by these EHS professionals regarding their visual inspection conduct. In comments similar to those reported by Woodcock, (2014) and Gallwey, (1998a) it was clear that these professionals did not see any problems with their idiosyncratic visual inspection conduct. This demonstrated that using a qualitative research approach to determine visual inspection performance by EHS professionals during visual inspections would be subject to a great deal of bias. This was qualitatively evidenced by responses that included “I don’t miss any hazards during my inspections”. Furthermore, at this early stage the complexities surrounding visual inspection conduct soon led to the realisation that it would require a credible level of inquiry, which best fitted a PhD level of research.

4.1 Epistemological Philosophy

And so began a period of reflection on just what evidence existed, and what data was needed to support the lead investigators long held confidence in systematic visual search (SVS). This led to a consideration of what epistemological philosophy could be used to generate reliable and valid data to inform visual inspection inquiry. Epistemology is defined by Colman, (2009) as “the theory of knowledge and especially what counts as knowledge”. It is further explained by Petty, et al (2012a) that epistemology involves itself with the overall strategies chosen as to how data is generated. There are two broad directions that epistemological philosophy can take either; positivist or interpretivist (Petty, et al (2012a). Positivism is defined by Colman, (2009) as the belief that “observation and experiment are the only sources of substantial knowledge”. The positivist epistemology (also known as scientific), would require an experimental design based on the exact control of the variables under analysis (Petty et al, 2012a). This in turn requires laboratory conditions and control over a small number of variables to generate data with a high level of reliability and validity. The type of data generated by this approach is largely quantitative.

Post positivism is further explained by Petty, et al (2012a) as also using an experimental approach for the manipulation of a single variable to explain any fundamental effects. In addition, it also accepts the limitations of a classic laboratory environment when investigating natural phenomena such as human behaviour. Therefore the manipulation of a single variable, for example the use of SVS, can be implemented within a post positive epistemology but without the narrow confines of an actual laboratory and precise measuring equipment. The post positivist tradition therefore allows for the use of functioning kitchens and an actual light aircraft to represent the laboratory. It also allowed the lead investigator to use lined or blank paper to be used as scripts, so that the lead investigator could subsequently record the percentage of hazards written down by participants in place of precise measurement. In the post positivist epistemological tradition, there can be a relaxing on the strict definition of experiment and measurement to accommodate situations where a laboratory environment is too restrictive for the inquiry of natural phenomena (Petty, et al, 2012a). However post positivism still retains the central tenet of the overall positivist approach, that of isolating one or more variables at a time and measuring the effect of any manipulation of those variables.

In contrast, the interpretivist direction, also known as constructivist or naturalistic (Petty et al, 2012a) would require a greater emphasis on the social world that visual inspectors inhabit and within which, their visual inspections are conducted. Interpretivist epistemology acknowledges that within the social world, the generation of a qualitative data set allows for a greater understanding of social reality, when that reality cannot be measured by precision instrumentation.

It was clear at this early stage that to comment with credibility on current and proposed professional visual inspection performance (VIP), reliable and valid evidence was needed. This was required to counter the inevitable scepticism that would be forthcoming from EHS professionals when making the assertion that their visual inspections could be flawed, or could possibly be improved upon. This was especially the case if SVS was demonstrated to be beneficial when compared to current custom and practice, as it could be seen as undermining previous risk assessments conducted by professionals. It was also clear that an ethical dilemma would be created if any recommendation for the use of SVS was continuing to be made by the lead investigator, without any supporting evidence. And so the first decision that needed to be reached by the lead investigator was the choice between a positivist or interpretivist approach. This choice is detailed in the next section.

4.1.1 The Decision to use a Positivist Against Interpretivist Approach

The choice of a positivist or interpretivist approach has a direct effect on the ontological status of the data generated. Ontology is defined by Colman, (2009) as the "study of the nature of being or existence". He goes on to explain ontology as concerned with the "distinction between reality and appearance, and whether mathematical realities exist outside people's minds". Examples of such mathematical realities include interval ratio data such as, temperature or weight and are viewed ontologically, as quantitative. Whereas narrative descriptions of people's reality such as how happy or sad they feel, are generally but not exclusively viewed ontologically as qualitative (Colman, 2009).

The research hypothesis states that; when looking for observable workplace hazards, the use of SVS will result in different visual inspection performance rates when compared to current practice. To recap from section 2.2.6, current visual inspection custom and practice is taken as the EHS professional walking around the site under analysis looking for, and recording, observable hazards. Therefore, when contemplating how to reject or not reject a future null hypothesis, the key verb to consider here is different. This verb responds very well to quantitative interval ratio data. It allows a degree of sensitivity to distinguish between any changes in percentage hazards observed, as a result of the SVS intervention in a statistically significant manner. The principal variable selected to be used to reject or not reject the research hypothesis was mean percentage hazards observed and referred to throughout this thesis as visual inspection performance or VIP. This is an interval ratio measure, which if generated with a high degree of reliability and validity would be very well placed to reject or not reject the null hypothesis. Therefore the research hypothesis, dictated in large part, the requirement for post positivist and thereby quantitative epistemological philosophy.

4.1.2 Experimental Design

The choices of epistemological philosophies to be adopted were influenced by the visual search literature, the research methods available, the pilot studies conducted, the research interests of the lead researcher and his practical limitations. However the overriding consideration was the requirement for a high level of reliability and validity for any data generated to evidence VIP. Therefore, the next step in the thesis was to select an actual experimental design and specifically the type of study. Study methods can be categorised into retrospective and prospective (Colman, 2009). A retrospective study uses historic or existing data whilst a prospective study looks to generate new data to answer a potential research hypothesis. An advantage of post positivist prospective studies is that they can be used to measure the effects of planned interventions such as the use of SVS on VIP. The research hypothesis therefore dictated that the best study would be prospective, of which there are two main categories; pre and post intervention studies and randomised controlled trials. The experimental design required a decision as to which type of prospective study to use and the following section will explain the decision to use a randomised controlled trial methodology.

4.1.3 The Randomised Controlled Trial Methodology

Of the research methods available for prospective studies, randomised controlled trial are often regarded as the best fieldwork methodology, and second only to controlled laboratory experiments conducted in the positivist tradition (Aronson, 2012; Cresswell, 2013). In the social sciences, randomised controlled trials have been described as the strongest research methodology available. (Aronson, 2013; Colman, 2009; Moher, 2010; Schulz et al, 2010). These scholars all state that the randomised controlled trial methodology when designed and conducted appropriately, represents the “gold standard” in research. Randomised controlled trial have a long and distinguished scientific history. The use of randomisation was first advocated by the statistician and geneticist Ronald Fisher in 1926 (Colman, 2009). The first randomised controlled trial was conducted in 1948 by Sir Austin Hill who researched the therapeutic efficacy of the antibiotic, Streptomycin (Colman, 2009).

Aronson, (2012) argues that the strength of the randomised controlled trial methodology lies in its robust ability to statistically evaluate the effects of any interventions, making it very appropriate for evaluating systematic visual search (SVS). It does this by randomising the assignment or selection of participants into comparable groups that do not receive the intervention (control) or do receive the intervention (experimental). If an appropriate random allocation method is used then comparability should be achieved resulting in participants in each treatment group being as similar as possible in terms of their inherent characteristics such as age experience or motivation. Aronson, (2012) further explains the advantages that the random assignment of participants to specific treatment groups bestows on the experiment. To explain, when random assignment of participants into treatment groups is conducted appropriately, any interventions are theoretically classified as being as free as possible from any selection bias. This in turn, allows causality again as far as possible of any effects observed in the experimental group and not in the control group, to be attributed to the intervention applied. Although the intention of randomisation is to achieve comparable control and experimental groups, it may not always achieve this. It is therefore necessary to review treatment groups to ensure as far as possible that these groups are constituted with similar participants, and are thereby capable of being appropriately compared (Aronson, 2012).

Although the randomised controlled trial approach is rooted in the positivist tradition (Petty et al, 2012b; Tabachnick & Fidell, 2014), the methodological epistemology behind the experimental design in this thesis is better described as being post positivist (Petty et al, 2012a). This is because bias from internal validity implies a measure of subjectivity regarding the complete isolation of controlled variables in the data elicited. This issue will be fully detailed in section 6.2.

In summary, the research hypothesis was to evaluate the effect of a systematic visual search intervention compared to visual inspection conducted according to custom and practice. Randomised controlled trials generate a quality of data second only to controlled laboratory experiments by being free from unevenly distributed participant characteristics. Their appropriateness for evaluating interventions is also well established. Given that controlled laboratory experiments would not have been feasible and could not provide ecologically valid conditions, the quality of data and appropriateness of the randomised controlled trial methodology were the overriding factors for its choice.

4.1.4 Further Reasons to not use an Intervention Study

The next best alternative to the randomised controlled trial would have been a pre and post intervention study. The primary advantages to using this epistemological methodology over randomised controlled trials would have been the need for less participants to achieve statistical power. A pre and post intervention study would only have required 64 participants in total, rather than a minimum of 128 for the eventual kitchen study (using an online statistical power calculator by Faul et al, 2007). The procedure for a pre and post intervention study would have involved these 64 participants being directed to carry out the observational task, prior to receiving instruction in SVS and then repeating the same observational task. There were additional logistical factors that also mitigated against a pre and post intervention study. These will now be detailed.

A pre and post intervention study would have been difficult to organise. This would have been due to the complexity of procedures involved in ensuring each participant correctly engaged in one visual inspection, followed by a training session, followed by exactly the same inspection task. Setting up a kitchen to take 64 pre-intervention participants would not have been possible due to the size of the actual kitchens. Instead numerous kitchens would have to be used. This would give rise to the attendant difficulty of invigilating all the inspections and thereby having to find sufficient personnel to assist the lead investigator, as well as recruit additional gatekeepers. Furthermore, if participants were allocated separate times for their inspection tasks, it would have been very difficult to prevent contact between treatment groups and thereby detrimentally affected the data generated.

In addition, ensuring that all the kitchens used were comparable would have been difficult to achieve in practice and bias could easily have arisen. The kitchens available to the lead investigator were fully functioning teaching kitchens and conditions within them would have been difficult to keep the same. Finally there would have been the issue of participants becoming suspicious of the methodology and trying to second guess, in particular, the post intervention observational task. As pointed out by Aronson, (2013) participants second guessing the motives of an intervention is a particular difficulty. Given these drawbacks, a pre and post intervention study design was rejected.

4.1.5 Ecological Validity

Having decided on using randomised controlled trials, the degree to which the trials were to reflect “real world” scenarios encountered by EHS professionals required consideration. Given the possibility of generalising the results to the wider EHS community, it was decided to incorporate as high a degree of ecological validity or realism, as possible. Ecological validity is defined by Colman, (2009) as “the confidence with which the conclusions of an empirical investigation can be generalised to naturally occurring situations in which the phenomena under investigation occurs”. In effect, the two thesis studies could be described as ecologically valid if they simulated real world visual inspections to a high degree. As a result, the decision was made to use the visual inspection of fully functioning industry standard kitchens, and an actual light aircraft to simulate “real world” visual inspection tasks. In short the kitchen and aviation studies needed to be as representative of the observational tasks that visual inspectors in these sectors routinely encounter. At the same time, this simulation of “real world” scenarios needed to ensure optimal control over the variables being investigated. In this way a high degree of reliability and validity in terms of evaluating the effect from any intervention was potentially possible.

The corollary of incorporating as much ecological validity as possible, meant that the creation of realistic simulated visual inspection tasks representative of “real world” workplaces was required. As it was logistically not possible to recreate randomised controlled trials for recruited participants in the EHS community, it was decided to create a small number of “real world” workplaces within the control of the lead investigator that would represent the wider workplace environment.

Hence in the kitchen study, the use of industrial grade kitchens from the lead investigators organisation which typified and reflected high quality catering facilities found worldwide. This also applied to the second study, the aviation maintenance inspection task whereby an actual light aircraft, located at the participating airfield was the subject under analysis. These simulated workplaces were designed to test and evaluate VIP by participants who were in large part professional visual inspectors themselves, or had aspirations to become so. In effect the participants were being asked to do what they were trained to do or did on a daily basis within a “real world” or ecologically valid settings; that of conducting visual inspections to see if their workplaces or equipment were safe to use.

In addition the participants in the two thesis studies could also be described as “ecologically valid” The kitchen study participants (N=211) were, safety and hygiene professionals with experience of visual inspections, or were studying to become so. The aviation study participants (N=26) were all very experienced visual inspectors. A full discussion of the participant characteristics is given in Section 4.2.8 and 4.3.5. Other than general bias issues, (see section 6.2) the nature of the participants recruited demonstrated that they were relevant, suitable for the intervention and did not present any apparent limitations for the two studies.

A further reasoning to consider a high degree of ecological validity was because the visual search literature calls for an improvement in “real world” visual search inquiry (Biggs & Mitroff, 2013b; Mitroff et al, 2015; Nakayama & Martini, 2011; Wolfe et al, 2015). Supporting a general call for an increase in “real world” interventions, particularly in the applied health sciences field. Zwarenstein, (2008) and Gaglio et al, (2014) both state that the demand for “real world” pragmatic randomised controlled trials is increasing. This is due to their inherent applied nature or usefulness when randomised controlled trials are used for research. Woodcock, (2014) has also called for more research into real world VIP although her study used an interpretivist epistemology.

The adoption of a high level of ecological validity also gave rise to additional potential advantages concerning generalisability. For example, if the intervention were found to be beneficial, then generalising the thesis results to the wider EHS community would become possible. Furthermore, previous findings by (Hrymak et al, 2015) had shown that EHS professionals did not have any standard visual inspection conduct. If it was found to be beneficial, SVS could potentially offer procedural standardisation for visual inspection conduct within the EHS community.

4.1.6 The Level of Pragmatism in the Randomised Controlled Trials

As this thesis consisted of randomised controlled trials, the consensus CONSORT approach for the reporting of such trials was followed as recommended by Schulz et al, (2010) and Moher, (2010). These scholars recommend that to correctly report a randomised controlled trial, there must be a clear methodology and findings presented together with a description of the level of pragmatism used in the study. The use of the CONSORT approach to reporting randomised controlled trials is recommended so that scholars can assess reliability and validity by using criteria common to healthcare related trials (Moher, 2010). CONSORT has consensus in the healthcare field and is advocated for example, by the British Medical Journal when reporting clinical trials.

The degree of pragmatism within any randomised controlled trial can be explained by using a spectrum analogy (Gaglio et al, 2014; Zwarenstein, et al, 2008). One end of this spectrum is characterised by the degree of exploration within the study, which seeks an explanation for the subject under analysis. The other end of this spectrum is characterised by pragmatism which seeks to evaluate the efficacy of the study to the “real world”. In other words, this pragmatism is a measure of how useful the findings of the study are on a spectrum that ranges from scientific knowledge to professional practice. It is also reported that there is a continuum between each end of this spectrum and where a particular study is located can be subjectively described (Gaglio et al, 2014; Zwarenstein, et al, 2008). These scholars dichotomise this distinction into efficacy, (can the intervention work and provide explanation), and efficiency (does the intervention work in practice and thereby reflect pragmatism). The purpose of a pragmatic randomised controlled trial within health care is to inform decisions about professional practice (Gaglio et al, 2014; Zwarenstein, et al, 2008). In summary, The randomised controlled trials conducted for this thesis can be described as located towards the pragmatic end of this spectrum.

4.1.7 Lessons From Two Pilot Studies Conducted Prior to the Trials

Having decided on an outline experimental design, it was time to see if this design could be translated into “real word” visual inspections conducted under ecologically valid conditions. So the next step was to pilot the experimental design. In this regard, two kitchens located within the lead investigator’s organisation were manipulated to ensure a known number of hazards, and participants directed to conduct visual inspections. The set-up of these kitchens and participant trial procedures are described in section 4.2.1.

Two pilot trials were conducted in November 2014 to assess the suitability of the experimental design. The first pilot did not involve randomisation due to an assignment process which simply did not work. Envelopes with sheets of paper numbered either 1 or 2 were attempted to be handed to participants. But these were opened early or swapped and the process was abandoned. The second pilot was conducted on a randomised controlled trial basis with the lead investigator assigning a 1 or 2 to each participant consecutively. This allocation method is accepted as being a source of bias and is detailed in section 6.2.1. Furthermore, the decision by the lead investigator to allow participants a ten minute period at the end of the inspection task to sit down and complete their writing, was problematical. It is custom and practice for the lead investigator and recommended by Hollis, (2000) to allow this period for reflection and final write up. Unfortunately, this ten minute period resulted in plagiarism being witnessed by the lead investigator and gatekeeper. Therefore the results from these two pilots were discarded from any analysis. Further lessons also emerged from these pilots. The number of hazard present in the observational tasks seemed to be about right for the inspection tasks set. The lead investigator timed himself at about 20 minutes to conduct a visual inspection for the two kitchens used. Participants were in the main, engaged with the observational task set although some control participants were noted finishing early which seemed to justify the 30 minute period chosen. The demand on the kitchens meant that there was only a short time available to use the kitchen. A final lesson was utility in familiarising the lead investigator with how to set up the kitchens and ensure the smooth running of subsequent trials.

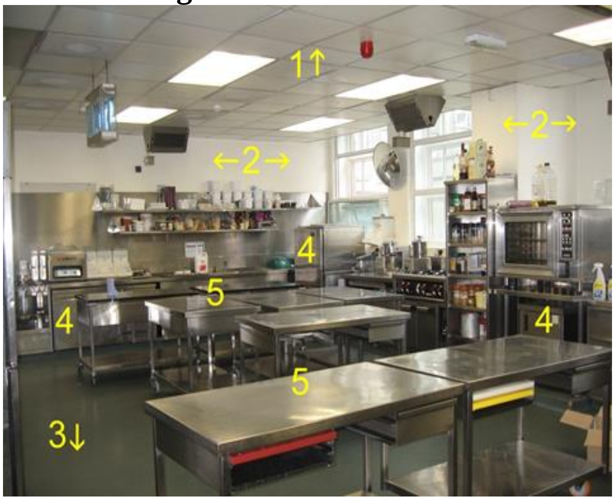
4.1.8 Experimental Design Summary

The experimental design chosen involved a post positivist randomised controlled trial epistemology, within a pragmatic and ecologically valid setting. Adopting a strict positivist epistemology with explicitly controlled laboratory conditions, would mitigate against extrapolating any findings to real world scenarios. In short, the increased variable control afforded by a strict positivist approach that laboratory conditions create, would have been at the expense of “real world” applicability. Internal validity and bias as a result of the ecologically valid experimental design data is acknowledged and accepted. However the greater potential “real world” use of SVS behaviour during inspections, was considered of greater value when compared to far simpler experimental designs allowing greater control over a smaller number of variables. This decision was further justified by the two pilot study results which demonstrated the utility of the chosen experimental design in generating reliable and valid data.

4.2 The Kitchen Study Method

The next section will detail an overview of the kitchen study method. It will include how participants were randomly allocated into treatment groups, and how experimental participants were trained in the SVS method. It will also detail how the observation of hazards were measured. To recap from Chapter 1 (Introduction), systematic visual search (SVS) is a method by which observers can conduct an exhaustive visual search of pre-selected areas using a consistent eye scanning pattern. It consists of two steps, the first being to select an element of a room for subsequent search. Experimental group participants in this study were thereby instructed that the strict order of selection was to be; ceiling first, followed by each wall, then the floor, then all storage items and finally equipment. A graphical representation of a typical kitchen and order of selection can be seen in Figure 8 below.

Figure 7 Selection Order for the Systematic Visual Search Method



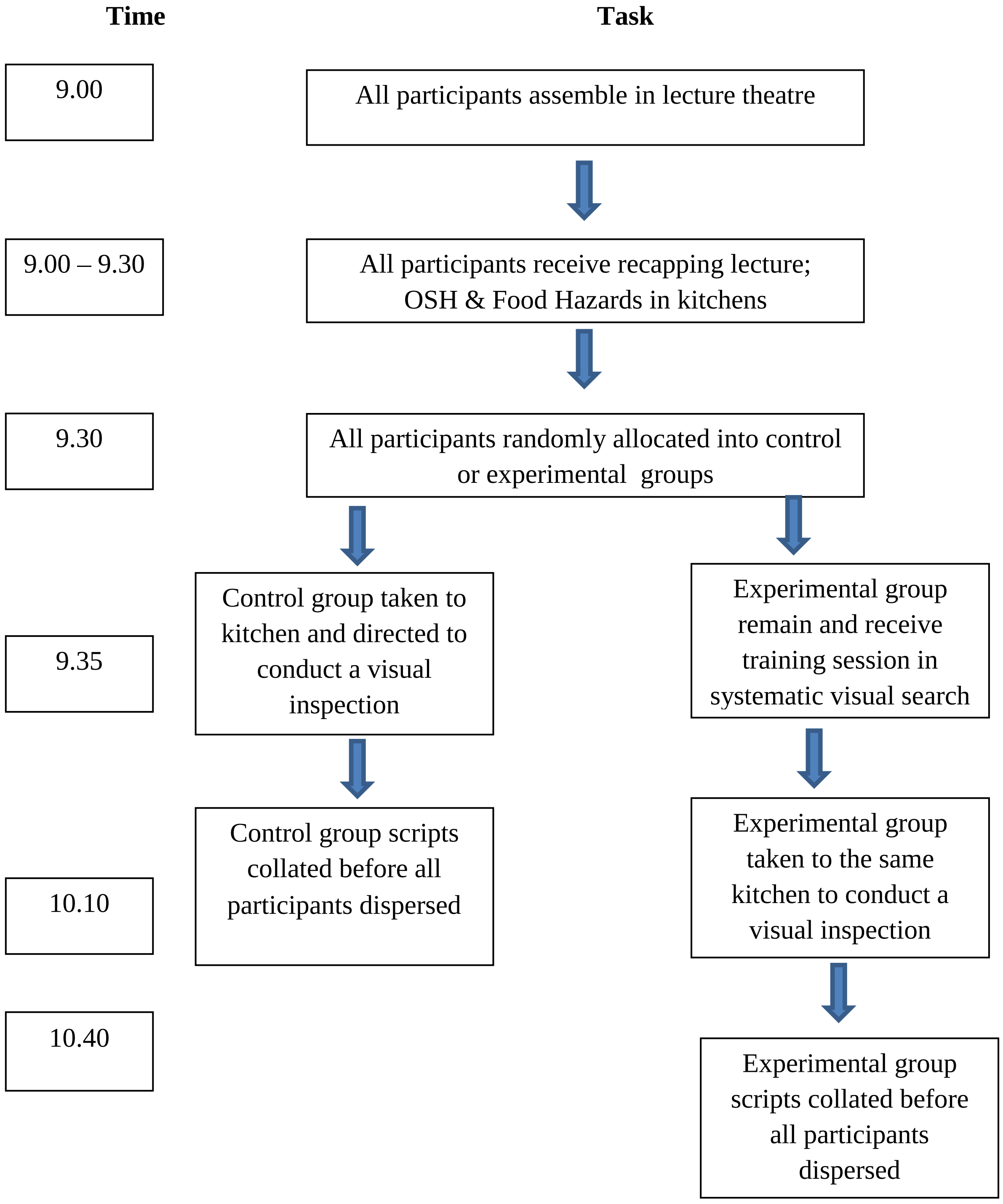
Once a room element (for example, the ceiling or a particular wall) was selected, it was subjected to the SVS method which uses an eye scanning pattern intended to ensure all areas under analysis are exhaustively searched. Again to recap from Figure 1, the eye scanning pattern used was the "reading a book" pattern. The next section will detail the procedures used in the randomised controlled trials for the kitchen study.

4.2.1 General Trial Procedures

Once all participants in a particular trial were assembled in a lecture theatre, a general recapping lecture was presented on common safety and food related hazards in kitchens. All subsequent hazards that participants encountered were included in this recapping lecture. The exact lecture notes and illustrative photographs used are shown in Appendix 1. Participants were then randomly allocated into control and experimental groups. A full discussion on the comparability of treatment group participants is given in sections 5.1 5.4.1 and 6.2.1 , where it is argued that the random allocation method resulted to a large extent in evenly matched participants and that there was no evidence that control or experimental groups were significantly different in any way. Together with the large number of participants, it is unlikely that any treatment group was different or better at the visual inspection task required.

Control participants were taken to a kitchen and directed to conduct a 30 minute visual inspection as per their normal custom and practice. Experimental participants remained in the lecture theatre and received a 30 minute training session on SVS (see Appendix 2 for the syllabus and section 4.2.3 for a precise explanation of the method). After this training in SVS, these experimental participants were then taken to the same kitchen as their control group colleagues and directed to conduct a 30 minute visual inspection using the SVS method. At the end of each visual inspection conducted, the written records produced were collected by the gatekeeper for analysis by the lead investigator. This procedure will now be described in detail beginning with a flow diagram for each trial procedure as shown in Figure 8 below;

Figure 8 Trial Procedure



4.2.2 Method of Random Allocation Sequence to Treatment Groups

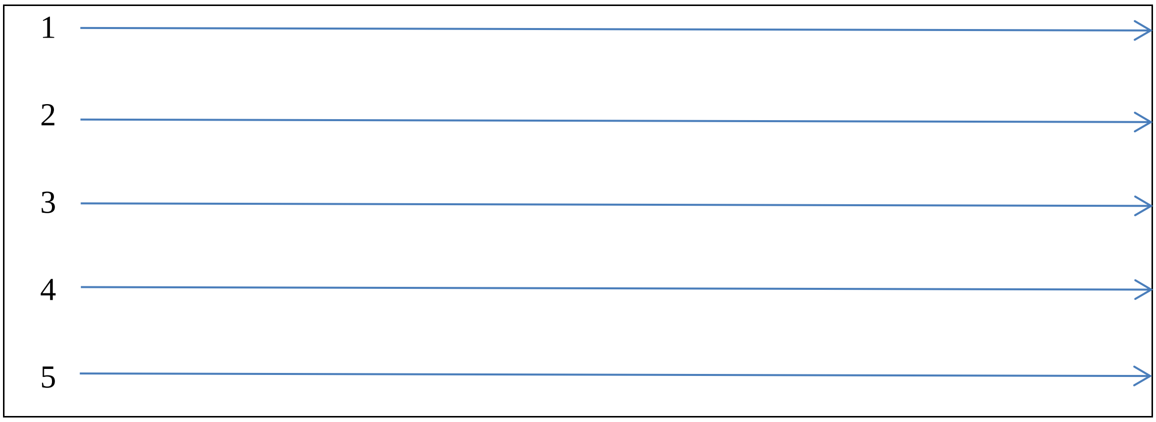
The sequence used in allocating participants to control and experimental groups in both studies was as follows. Control group participants were to be assigned the number 1 and experimental group participants were to be assigned the number 2. The lead investigator chose the first participant seated at the right hand side of the bottom lecture theatre row. He pointed at this participant and verbally assigned them the number 1. The participant next to them was assigned the number 2 and this continued consecutively in a "snakes and ladders" pattern until all participants had been sequentially assigned. This assigning was done at the end of the re-capping lecture. Having created the control group, the lead investigator informed these participants that they were to be taken by the gatekeeper present, to the selected kitchen. Here they were required to visually inspect and write down all the food and safety hazards they observed within thirty minutes on A4 paper. The gatekeeper also acted as an invigilator during this thirty minute inspection and collected all the scripts at the end.

4.2.3 A Practical Description of Systematic Visual Search

The behavioural visual search algorithm that constitutes SVS, can now be described in detail. SVS will be practically explained by outlining how a visual inspection of a room would be conducted using this method. For this exercise, readers are asked to conduct a visual inspection of the room they are currently sitting in. The first step with SVS is to select one element of the room. In construction engineering jargon, elements are integral constructional objects that separate different spaces. Building elements will include, the ceiling, each wall, the floor and doors. The element that will be used to demonstrate SVS to readers is a wall in the room they are currently occupying.

Accordingly, readers are asked to look directly at the wall in front of them, and asked to fixate on (stare at) the top left hand corner of the wall. Then, with an eye movement that scans to the right, the reader is directed to look along the length of the wall until the right hand side of the wall is reached. Readers eyes should then directed back to the left hand edge of the wall, but underneath the area already observed again scanning to the right until the left hand edge of the wall is reached. And so on until the entire wall has been exhaustively searched. A further graphical description is given below in Figure 9 below. Readers are asked to imagine the rectangle below is the outline of the wall in front of them. Readers should begin by fixating on the number one figure located in the top left corner, and then follow the line of the arrow until the arrow head is reached. At the arrow head the reader then relocates his or her gaze to line number 2, and so on until the entire rectangle is observed exhaustively.

Figure 9 Describing the "Reading a Book" Eye Scanning Pattern



This behavioural visual search algorithm is methodologically in stark contrast to top down approaches such as checklists which can lead to unpredictable eye scanning patterns of the inspection space. In contrast, SVS is designed to allow the eye scanning pattern to ensure all areas are visually searched. In this way, objects of interest should (in theory) be fixated on and thereby observed when eye movements arrive at the objects location.

4.2.4 Delivering the Training for the Systematic Visual Search Method

A 30 minute training session in SVS was delivered in all the trials by the lead investigator in a lecture theatre. The educational pedagogy used was a PowerPoint presentation for the first 15 minutes explaining SVS, and that this method was to be used for the forthcoming kitchen visual inspection (see Appendix 2 for the syllabus and Appendix 3 for the slides used). This was followed by a demonstration of SVS by the lead investigator, before inviting participants to conduct a "practice" SVS themselves, in the actual lecture theatre. Experimental group participants were first shown the order in which to select room elements for subsequent SVS use being the ceiling, north wall, east wall, south wall, west wall, floor, storage, and finally equipment (see Appendix 4). Participants were informed that within the kitchens, each wall had a sheet of paper stuck high up, with N, S, E & W to facilitate the selection order for observing walls.

The use of SVS began with the lecture theatre ceiling. This element was observed in its entirety by the lead investigator demonstrating the “reading a book” eye scanning pattern with any hazards pointed out to participants. A wall was then searched by the lead investigator again using the same “reading a book” eye scanning pattern and pointing out any hazards. Once SVS was demonstrated on the ceiling, a wall and the floor, participants were invited to practice SVS from their seats using a wall. On discovering a hazard, such as a leaking pipe, broken glazing or unhygienic conditions participants were instructed to stop observing and to write the hazard down. Once this observed hazard was written down, participants were instructed to return to the point at which they had observed the hazard, to resume their visual search. The importance of eye scanning the entire element using the “reading a book” pattern was stressed on a number of occasions. To ensure foveal resolution, experimental group participants were also shown by the lead investigator to walk the area they were visually searching. This was to bring themselves into close contact with the area under visual search analysis.

Where visual search was not possible due to obfuscation or no direct line of sight, experimental group participants were instructed to manipulate the environment to see if visual search could be improved. This was demonstrated by the researcher in the lecture theatre by moving curtains aside, trying to pull anything that had a handle or was capable of being opened pushed or moved. This point was particularly reinforced with the floor element. Here the lead investigator repeatedly made the point that all areas of the floor were to be visually searched by walking over all areas. It was also stressed and that in some instances, it was necessary to get down on “all fours” to fully observe the floor in its entirety and especially under any floor mounted equipment. Finally experimental group participants were instructed that only hazards were to be written down. Instances of no access were not to be recorded. This was directed in order to maximise the inspection time available to experimental group participants.

This training session was supported by a nine page instructional document (see Appendix 4) given to participants in order to further facilitate the use of SVS during the actual kitchen visual inspection. This handout was distributed to experimental participants just after the lead investigator had demonstrated SVS on the ceiling, a wall and the floor. The first page of this document included a graphical representation of the order in which to select room elements for subsequent SVS. An additional eight sheets of paper were attached to this document, each of which was separately and consecutively labelled as follows; ceiling, north wall, east wall, south wall, west wall, floor, storage, and finally equipment (see Appendix 4). This order again facilitated experimental participants to follow the order given and to use SVS on each room element. In addition and for experimental design purposes, the first page of this documentation requested demographic data; gender, job title and number of years working.

On expiry of the 30 minute training session, the lead investigator escorted the experimental group participants to the same kitchen visually inspected by their control group colleagues some thirty minutes previously. On arrival at this kitchen, the experimental group participants were thereby invigilated for the thirty minutes by the lead investigator and gatekeeper. At the end of this period, all documents were collected and given to the gatekeeper as per ethical approval procedures. The gatekeeper ensured participants could not be identified from the A4 pages or documents before returning the data to the lead investigator for subsequent analysis.

In concluding this explanation of the training method used for systematic visual search, the following should be noted. The thesis intervention consisted of training in systematic visual search, the PowerPoint presentation as well as the nine page handout used to write down observed hazards by experimental group participants. Therefore and throughout this thesis, the intervention is taken to constitutes these three separate components of the training in SVS itself, the lecture slides presented and nine page handout used.

4.2.5 The Creation of the Simulated Kitchen Inspection Task

A total of five fully functioning, industry standard kitchens, located within the Dublin Institute of Technology Cathal Brugha St campus, were selected for use. These kitchens are used for the education and training of professional catering staff and chefs, leading up to degree level awards. These kitchens were designed and constructed in order to provide high quality, “fine dining” cuisine. Kitchens area and volumes are shown in Table 11 below.

Table 9 Dimensional Description of the Kitchens Used

Kitchen	**Area*m^2**	**Volume*m^3**
1	97	278
2	113	317
3	73	204
4	75	217
5	117	294
	M=95;	M = 262

4.2.6 A Description of Hazards in the Kitchen Study

The hazards within the kitchens, served to evaluate the effect of SVS as an intervention and therefore needed careful detailing. A full listing of the hazards used across all the trials is included in Appendix 6. Here a generic description of the hazards within the kitchens will now be presented. There were two types of hazards encountered by participants, those that were already present in the kitchens and those “planted” by the lead investigator for experimental reasons.

The existing hazards mainly included examples of substandard cleaning and disrepair but did not represent the wide range of hazards that could be typically found within kitchens. The main reason for planting additional hazards therefore, was to ensure participants had an observational task with a high degree of ecological validity. Hence it was necessary to ensure a range of different hazards were contained within the kitchens that would reflect “real world” conditions. These planted hazards were selected by the lead investigator, drawing on his 30 years of experience of kitchen inspections. Although a subjective judgement, the existing and planted hazards together represent commonly encountered kitchen based hazards. Planted hazards included food left out in the open, cross contamination in fridges and freezers, cigarette materials in ash trays, covered smoke detectors, and moved ceiling tiles, (commonly encountered when routine maintenance is carried out) fire extinguishers propping open fire doors and overloaded sockets. Where hazards were ambiguous or needed specialist knowledge, for example machinery guarding requirements for industrial grade food mixers, they were removed. The kitchens were left so that the mean number of hazards per trial was kept as uniform as possible with M=36.17 (SD= 4.04).

The hazards were categorised into occupational safety and health (OSH) hazards and food hazards. OSH hazards included pictogram labelled bottles showing toxic or flammable properties, bare wires, covered smoke detectors, or moved fire extinguishers. Food hazards were categorised as capable of causing food poisoning, pathogenic cross contamination or constituting poor hygiene conditions including; food left in open, food on floor, unhygienic walls floors or windows, raw and cooked food stored close together in a fridge or freezer.

Open view hazards were regarded as those that were capable of being observed, by direct line of sight by the participant. Obfuscated view hazards were regarded as those that required participants to manipulate the immediate environment or positioning themselves for a better view for example, opening a fridge door to reveal cross contamination, looking behind equipment at wall surfaces or kneeling down to look under equipment. High risk hazards were considered to have a high potential of human harm compared to other hazards and were; covered or interfered smoke detectors, toxic and flammable liquids, fire doors left open, fire extinguishers moved from position, bare electrical wires, food left out in the open, smokers materials, raw and cooked food in close contact, and mouse traps.

4.2.7 Ascertaining the Pre-Trial Number of Existing Hazards in Kitchens

Prior to each trial, and usually an hour or so before participants were due to assemble in the lecture theatre, the lead investigator conducted a visual inspection of the kitchen using SVS. He thereby noted all the hazards visually observed. In addition he manipulated the kitchen by "planting" a set number of hazards for subsequent observation by participants. The lead investigators colleague, the study's gatekeeper, then joined him in the kitchen and also conducted her own pre-trial inspection. The gatekeeper in discussion with the lead investigator then confirmed the list of the "known" number of hazards prior to each trial. The kitchen was then locked by the gatekeeper to prevent unauthorised access.

4.2.8 A Description of the Kitchen Study Participants & Trials

A description of the participants will be presented here and will include descriptive statistics of the trials, recruitment details and deselected participants. Table 12 below, details the kitchen trials conducted.

Table 10 Procedural Details of the Kitchen Trials Conducted

Trial No.	**Date**	**Kitchen No.**	**Participant work status**	**Total N**	**N Control**	**N Experimental**	**N Control Excluded**	**N Experimental Excluded**	**N Hazards Present**
1	Dec'14	3	FT* Students	11	6 (3,3)	5 (4,1)	0	0	27
2	Dec'14	3	Professional	13	6 (2,4)	7 (3,4)	0	0	32
3	Dec'14	2	FT Students	12	6 (4,2)	6 (3,3)	0	0	43
4	Jan'15	1	Professionals	16	7 (6,1)	9 (2,7)	0	1	37
5	Jan'15	1	Professionals	9	6 (3,3)	3 (0,3)	0	2	38
6	Feb'15	5	FT Students	20	9 (5,4)	11 (4,7)	2	3	34
7	Feb'15	1	PT* Students	19	9 (2,7)	10 (5,5)	1	1	40
8	Feb'15	4	PT Students	19	11 (4,7)	8 (3,5)	0	2	36
9	Feb'15	1	PT Students	11	6 (5,1)	5 (4,1)	0	0	35
10	Feb'15	1	PT Students	22	11 (4,7)	11 (4,7)	0	0	38
11	Mar'15	1	FT Students	24	12 (5,7)	12 (4,8)	0	0	38
12	Apr'15	4	FT Students	35	18 (1,17)	17 (13,4)	0	1	36
Totals				**211**	**107**	**104**	**3**	**10**	**M=36.17** **SD=4.04**

FT* = Full Time, PT** = Part Time (figures in brackets show male and female numbers)

4.2.9 Participant Recruitment

A total of 224 participants were recruited for the kitchen inspection task. The number of participants subsequently deselected was 13 (see Table 12 above), leaving 211 active participants. Of these 211 participants, 154 were undergraduates from the Dublin Institute of Technology Cathal Brugha Street Campus. These participants were all studying for level 6, 7 or 8 undergraduate programmes where safety and hygiene inspection duties would be a daily event in their future careers. The relevant programmes all involve food safety or occupational safety (OSH) related programmes. Ethical approval anonymity requirements prevent the naming of these programmes, but they can all be described as offering professional career opportunities within the EHS community.

4.2.10 Participants Excluded from the Kitchen Study

The deselection of participants who do not adequately take part in randomised controlled trials studies is an accepted practice. (Zwarenstein, 2008; Gaglio et al, 2014). These authors point out that the correct reporting of randomised controlled trials requires any participant withdrawals to be reported. The reasoning behind the withdrawals also needs to be detailed as is the case below.

There were a small number of participants who were excluded from the kitchen study. The scripts from a total of thirteen participants, (three control group and ten experimental group) were not used. Those deselected represented 5.8% of the total participants recruited. Furthermore the results chapter will show that VIP between the control and experimental groups would not be significantly affected by the deselection process used here. Table 12 above shows which trials these deselected participants came from.

In these particular deselected cases, it was clear that there was a very poor motivational effort to appropriately participate. This lack of motivation manifested itself as a lack of sufficient writing down of observed hazards. The mean number of hazards written down in the control and experimental groups in the kitchen study was some thirty one and forty eight respectively (see Table 15). Where the total number of hazards written by participants failed to reach double figures, or it was clear the inspection task was not being taken seriously, the results were not used. In addition, these deselected participants could also have been viewed as statistical outliers, an additional reason for their exclusion (Field, 2013). However, why ten experimental participants did not show sufficient motivation compared to three control participants is difficult to speculate on. The main reason seems to be the nature of the interventional task. Experimental participants were directed to change their inspection behaviour which requires effort. If lack of motivation to participate was a factor, it was more likely to manifest itself within the group that required more effort being experimental participants. However this remains pure speculation and can also be explained by sample error amongst the participants.

4.3 The Aviation Study Method

On completion of the kitchen study, a partial analysis of the results were completed, which as will be seen in Chapter 4 (Results) began to show the potentially positive effects from the SVS intervention. On reflection at this stage, there was one area of bias that could not be easily eliminated and caused by the experimental design. It became apparent that positive results for the intervention could have been due to a motivational effect from the lead researcher on experimental group participants. It was clear that the only way to counter this, was to run a separate study whereby the lead investigator had no contact with participants. Hence the recruitment of aviation engineers for a second study.

4.3.1 General Trial Procedures

The procedures were very similar to the kitchen study. For brevity it is advantageous to detail the differences between the kitchen and aviation study rather than similarities. The main differences were that the lead investigator was not involved in the delivery of aviation study training session, or the manipulation of the light aircraft used. Furthermore, the number of participants was smaller with a total of 26 experienced aviation maintenance engineers. The aviation study was entirely administered by two senior aviation engineers recruited by the lead investigator. Both these senior aviation engineers were trained in the SVS method prior to the randomised controlled trial by the lead investigator. Due to anonymity conditions requested by the aviation maintenance organisation, it was not possible to collate as much background information as with the kitchen study. Therefore specific detail on the exact syllabus, pedagogy and dimensions of the light aircraft are not available.

4.3.2 The Creation of the Simulated Aviation Inspection Task

The visual inspection task was a light aircraft. The participants were all experienced aviation maintenance engineers. An example of the type of light aircraft used in shown below in Figure 10 below.

Figure 10 The Type of Light Aircraft Used

4.3.3 A Description of the Hazards in the Aviation Study

All the hazards used in the aviation study are listed in Appendix 7. As with the kitchen study the senior aviation engineers were asked to prepare the light aircraft with a range of existing and planted hazards (defects in their jargon) that would simulate an ecologically valid “real world” pre-flight inspection. The criteria for high importance, was decided upon by the senior aviation engineers involved. Their criteria applied to all variables was a dichotomous low or high importance rating. This decision was based on the opinion of these two senior engineers as to possibility of the hazards (defects) leading to negative in flight consequences, if not remediated.

4.3.4 Ascertaining Pre-Trial Number of Hazards on the Light Aircraft

Prior to the trial, the two senior aviation engineers carried out a joint inspection of the light aircraft and noted all the existing hazards (defects). These engineers were asked by the lead investigator, and agreed to use the systematic visual search method for their visual inspection. In addition they manipulated the light aircraft by “planting” a set number of hazards (defects) for subsequent observation by participants.

4.3.5 A Description of the Aviation Study Participants

Twenty six engineers were recruited and one randomised controlled trial was conducted. Groups of 4 engineers inspected the aircraft with the same thirty minute time frame. All educational interventions and supervision was conducted by the two senior aviation engineers. The control group used their own in house documentation to record all hazards. The senior engineers requested that this documentation would be kept confidential. However it can be reported that “standard aviation reporting” documentation was used by the control group participants. For the experimental group participants, the supporting documentation was designed by one of the senior engineers and was seen by the lead investigator to be broadly similar to the kitchen study documentation. The pre-trial number of hazards was recorded by the two senior engineers and was 29 in total. There were no deselected participants.

4.4 Ethical Approval

All pilot studies and fieldwork for this thesis was conducted after ethical approval was granted. A copy of the ethical approval is included in Appendix 7.

4.5 The Statistical Analysis Plan

This section will detail the mathematical treatment of the data generated. It will begin by detailing how the raw data were converted into scores and the variables created. It will then detail the descriptive and inferential statistical tests used.

4.5.1 Data Generation

For both the kitchen and aviation study, the data generation involved the lead investigator reading each hazard written down on the scripts. These hazards were then compared to the list of pre-identified hazards displayed in a Microsoft excel 2010 programme file. In this way each hazard the participant identified, and that matched the pre-existing list of hazards, was recorded. In addition gender, job title and number of years working were also recorded. Once all data was inputted into an Excel file it was transferred to an IBM version 21 SPSS programme file.

The first mathematical calculation carried out was to convert the number of hazards identified by participants into percentage scores of the overall number of hazards written down. This allowed comparative scores to be standardised. The mean percentage of hazards identified by participants became the de facto measure of visual inspection performance (VIP). This mathematical calculation has been used previously in the visual search literature for example by Albert et al, (2014 & 2017); Gallwey (1998a). Once the relevant data had been converted into percentages, then measures of central tendency, spread, effect size, means testing, correlation and regression values were able to be calculated.

Examples of the calculations used to generate data will now be detailed using four hazards from the kitchen study. These hazards are whether participants observed; a smoke detector covered by a lunchbox, a cooked pie left on a shelf, a fire door wedged open by a fire extinguisher and cross contamination in a fridge.

These examples are shown below in Table 13 below which has these four hazards treated as variables, together with three participant scores. An extended version of this table was created in Excel and SPSS. A figure of zero indicates the participant did not record the hazard during their inspection and a figure of 1 indicates that they did.

Table 11 Example of Data Generation

	Hazard Type	**Hazard Type**	**Hazard Type**	**Hazard Type**	**Measurement**
Partici-pant Number	**Smoke Detector Covered**	**Cooked Food Left in the Open**	**Wedged Fire Door**	**Cross Contamination in a Fridge**	**Percentage of Total Hazards Observed (VIP)**
1	1	1	1	1	100% (4 out of 4)
2	0	0	0	1	25% (1 out of 4)
3	0	0	0	0	0% (0 out of 4)
					* Overall Mean 41.67%
	33% observation (1out of 3)	33% observation (1out of 3)	33% observation (1out of 3)	66.67% observation (2 out of 3)	

The final column in Table 13 above represents the overall results standardised as percentages. The scores of each participant as well as the overall mean percentage of all hazards observed by all participants* is include here as 41.67%. This score; the overall mean percentage of all hazards observed by participants is phrased throughout this thesis as, visual inspection performance or VIP. This was done to highlight its title as reflective of an overall visual inspection measure. Similarly, individual hazards types had their observation rate measured so that the VIP for smoke detectors covered, is shown at the bottom of the first row, as 33%. Aggregating individual hazard results allowed constructs such as smoke detector scores, added to wedged fire door scores, to became a measure of occupational safety and health. The treatment here was to convert all frequency scores into percentage scores. This was required as the number of hazards per individual trials varied and using frequency scores would be mathematically incorrect. Over 22,000 data points from two studies were generated. This provided a large data set for potential analysis. How this data set was mathematically manipulated and treated will now be detailed.

4.5.2 The Extraneous Independent Variables Created

The use of SVS during inspections, was the independent variable created and experimentally manipulated as part of the randomised controlled trial methodology for both studies. The remaining extraneous independent variables created, and not experimentally controlled for the kitchen study are as follows; the kitchen inspected, gender, experience, work status, trial number, the number of hazards within the kitchen and finally kitchen area. For the aviation study, due to participant restrictions it was only possible to gather data on one extraneous variable, the number of years worked.

4.5.3 The Descriptive & Inferential Statistics Tests Used

Once the scores had been standardised into percentages then the results could be displayed for both control and experimental group participants. They are listed below with the descriptive and inferential statistics and tests used;

- Percentage observation of all hazards (visual inspection performance or VIP)
  - Histogram distributions, and bar charts
  - Means, Standard Deviations and 95% Confidence Intervals
- Percentage observation of category of hazards
  - Histogram distributions, and bar charts
  - Means, Standard Deviations and 95% Confidence Intervals
- Number of hazards written down by participants
  - Histogram distributions, and bar charts
  - Means, Standard Deviations and 95% Confidence Intervals
- Pearson’s Product Moment Correlations
- Regression Analysis
  - B, β and R^2
- Percentage and frequency differences between treatment groups
  - Independent t testing with a Bonferroni correction
  - ANOVA
  - Theta; η
  - χ^2
  - Cohen’s *d*

4.5.4 The Null & Alternative Hypothesis

Framing the research hypothesis in null terms (H_0) the statement is; when looking for observable workplace hazards, the use of SVS will ***not*** result in different visual inspection performance rates when compared to current practice.

The alternative hypothesis (H_1) is; when looking for observable workplace hazards, the use of SVS ***will*** result in different visual inspection performance rates when compared to current practice. A decision on which hypothesis to accept or reject is given in section 5.3.

4.5.5 Significance Testing for Differences Between Treatment Groups

In classic Null Hypothesis Significance Testing, significant differences between group means are ascertained by using an independent *t* test. In order to reject or not reject the research hypothesis, the significance value of a = .05 was assigned as recommended by scholars in this discipline (Field, 2013; Tabachnick & Fidell, 2014). Unless otherwise stated statistical significance is taken throughout this thesis as $p \leq .05$.

However, as there were many other variables recorded further *t* testing between treatment groups would inflate familywise error and therefore could not be recommended. Two further statistical tests were used in such cases, a Bonferroni corrected independent *t* test and ANOVA. A "Bonferonni" correction, as recommended by Field, (2013) was therefore applied to the remaining variables selected for comparison. The formula used was;

$$P_{\mathrm{crit}} = \frac{\mathrm{a}}{\mathrm{k}}$$

where a represents the *p* value used and k represents the number of cases or tests to be used. A Bonferonni correction was used as well as ANOVA for its robust nature and as a result, the significance value for the Kitchen study was; a = ≤ 0.0017 and for the aviation maintenance study a = ≤ 0.0016. The use of a Bonferroni correction has been described as robust by statistical scholars (for example by Field, 2013) and represents a conservative approach when concluding on differences between selected variables or treatment groups. The measure used to calculate effect size was Cohen's *d* and was calculated by;

$$d = \frac{X1 - X2}{s}$$

where

X1 = mean of experimental group

X2 = mean of control group

s = standard deviation of the control group

The effect size descriptors as recommended by Field, (2013) for Cohen's *d* are;

- Less than 0.2; small effect size
- 0.5; medium effect size
- Greater than 0.8, large effect size

4.5.6 Correlations & Regression

All variables selected for correlation used Pearson's product moment correlation co-efficient, *r*. Field, (2013) recommends the following effect size measures for *r*, and these values have been used throughout this thesis;

- *r* ±.1 represents a small effect or 1 % of the variance
- *r* ±.3 represents a medium effect or 9 % of the variance
- *r* ±.5 represents a large effect or 25 % of the variance

ANOVA was also used to see what effect extraneous variables had on visual inspection performance. In reporting ANOVA results where η^2 values are presented, effect sizes are taken to be the same as *r* (Field, 2013). In calculating effect sizes for ANOVA, η^2 was used with the following formula from Field, (2013);

$\eta^2 = SS_m/SS_t$ where SS_m is the model sum of squares and SS_t is the total sum of squares.

In terms of regression analysis, Tabachnick & Fidell, (2014) point out that in experimental designs that use randomised controlled trials, treatment groups can become separate populations for analysis. However this is the case only if the statistical tests show significant differences due to the intervention. As significant differences were found in treatment group means in this thesis, simple regression was used to measure the strength of the intervention. This approach has been used in the visual search literature. Biggs & Mitroff, (2013b) for example used regression analysis to compare professional and non-professional searchers on a similar visual inspection task.

4.5.7 Statistical Power Analysis

As commentators point out, the sample size used for any study is an important parameter (Cohen, 1992; Field, 2013; Tabachnick & Fidell, 2014). The sample size should be big enough to ensure that the probability of type 2 error rate (false negatives or missing an effect) is minimised. Statistical power refers to the probability of finding an effect within a population. According to Field, (2013) a consensus figure for this probability is .8 representing an 80% chance that any effect will be found. Sample size calculations are relatively complex but Field, (2013) advises on the use of internet based sample size calculators such as G*Power (Faul et al, 2017). Using this web based tool, the parameters used to calculate statistical power were; α = .05, and 1- β, = 0.8 or 80% power. The resultant sample size for the randomised control trial was 64 control and 64 experimental participants. The actual sample sized used was 107 control and 104 experimental participants which using the same programme returns a statistical power of .9 The study can be therefore described as being above the requisite level of statistical power to detect any significant findings.

4.5.8 Summary of the Methodology

A total of 237 participants were recruited for a Kitchen and aviation visual inspection study. For the kitchen study, 211 participants were recruited with 12 randomised controlled trials resulting in a statistical power of .9. For the aviation study 26 participants were recruited, but statistical power was not achieved. However, it should be noted that the main reason for the aviation study was to eliminate any lead investigator bias by removing him from this intervention, and investigate any further applicability. In total circa 22,000 data points were generated allowing a statistical evaluation of the effect of an SVS intervention, conducted under ecologically valid conditions.

CHAPTER 5 RESULTS

5.01 Introduction

This chapter will present results from the thesis field work stage. Using this data, it will argue that the randomised controlled trial methodology resulted in control and experimental participants of sufficient comparability allowing systematic visual search, to be robustly evaluated as the intervention. It will further argue that the results are able to support the alternative research hypothesis and state that; the use of SVS will result in different visual inspection performance rates when compared to current practice. In order of presentation, the kitchen study will be presented first, followed by the aviation study. For each study, a demographic comparison for treatment group participants will be detailed first before presenting descriptive and inferential statistical results. The chapter will finish with a comparative analysis between the two studies.

5.1 The Kitchen Study; Descriptive Results

This section will begin by presenting demographic data testing for the level of similarity between treatment group participants. It is important for randomised controlled trials, to show just how comparable the control and experimental groups are to each other by detailing their respective demographic data. Table 14 below presents this comparable data. This demonstrates that the intervention was applied to similar participants, and that one group was not biased in any way following random allocation.

5.1.1 Demographic Comparison Between Treatment Group Characteristics

Table 12 Demographic Comparison of Treatment Groups in the Kitchen Study

Categories	Control	Experimental	Total
N	**107**	**104**	**211**
N Males	**44**	**49**	**93**
N Females	**63**	**55**	**118**
N full time students	**51**	**51**	**102**
N full time students with no workplace experience	**24**	**23**	**47**
N full time students with at least 1 year of experience	**27**	**28**	**55**
N part time students	**37**	**34**	**71**
Mean years of experience of part time students	**M= 4.26 SD = 3.02**	**M= 7.39; SD = 5.74**	
N full time professionals	**19**	**19**	**38**
Mean years of experience of full time professionals	**M= 16.47; SD = 10.30**	**M= 13.79; SD = 10.50**	

What can be seen with the kitchen study participants in Table 14, is the closeness in demographic characteristics between control and experimental groups. They were similar enough in all categories, to warrant a valid comparison and ensure any effects could be attributed as far as possible to the intervention. In addition, and in order to strengthen the evidence for comparability a χ^2 test was used and found no significant differences between the ten categories in Table 14 above ($p \leq .05$). This test provided additional empirical confidence for ensuring the comparability of control and experimental participants.

5.1.2 Descriptive Results from the Kitchen study

The overall descriptive and statistical data from the kitchen study will be presented in Table 15 below. It should be noted that the control and experimental columns present the mean percentage hazards observed (VIP) for each category of hazard listed in the first column. Table 15 below, clearly demonstrates the beneficial effect of SVS when used to observe the hazards and constructs listed compared to the control group. The comparability of participants, and the randomised controlled trial methodology, strongly supports the theory that this beneficial effect on VIP, was caused by the use of SVS. This finding is further tested against null and alternative hypotheses in section 5.3.

In overall terms, the mean percentage of all kitchens hazards observed by SVS users was 16.67% more than their control group colleagues. This finding was significant ($p \leq .0017$). Furthermore, the effect size as measured by Cohen's *d* was large, demonstrating the strength of the intervention. In addition, Table 15 shows that for 15 out of the 23 hazards or constructs listed, the effect size was medium or large again demonstrating the strength of SVS on the observation of hazards. The effect of SVS on the observation of kitchen hazards can also be seen graphically in Figure 13.

Nevertheless, there was a wide range of VIP scores demonstrated. The control group participants ranged from a low of 10.51% for observing mouse traps, to a high of 55.84% with the observation of food packets left on the floor. The experimental group participants ranged from a low of 31.01% for the same mouse traps, to 87.98% observation of moved ceiling tiles. Clearly, using SVS was by far the largest predictor for improving VIP but some rates of observation were closer between treatment groups, a point that will be detailed in section 5.1.5.

Table 13 VIP Data by Specific Hazards & Constructs; Kitchen Study

Hazards Identified	**Control (VIP M&SD) [95% CI]**	**Experimental (VIP M&SD) [95% CI]**	**%age difference in M's**	**Cohen's *d***	***p*= ≤.0017***
Overall Mean	32.97 (9.03) [31.24- 34.70]	49.64 (10.89) [47.53- 51.76]	16.67	1.85	✓
Occupational Safety & Health related	39.42 (19.24) [35.73-43.11]	43.30 (18.59) [39.68-46.91]	3.38	0.2	x
Food related	30.16 (9.85) [30.16-28.28]	53.44 (13.48) [50.82-56.06]	23.28	2.36	✓
Open View	41.78 (11.90) [39.50-44.06]	54.32 (12.25) [51.94-56.71]	12.54	1.05	✓
Not in Open View	17.63 (10.88) [15.54-19.71	42.40 (17.46) [39.01-45.80]	24.77	2.28	✓
All High Risk	31.42 (14.06) [28.73-34.11]	41.65 (15.37) [38.66-44.64	10.22	0.73	✓
OSH High Risk	35.96 (24.27) [31.31-40.61]	45.11 (23.24) [40.59-49.63	9.15	0.38	x
Food High Risk	27.69 (15.76) [24.67-30.71	37.69 (21.25) [33.56-41.83]	10.01	0.64	✓
OSH Open View	37.90 (29.46) [32.25-43.54]	43.17 (26.43) [38.03-48.31	5.27	0.18	x
Food Open View	57.17 (38.99) [49.69-64.64]	61.37 (41.82) [53.24-69.51]	4.2	0.11	x
Food High Risk Not in Open View	20.30 (19.51) [16.56-24.03]	31.70 (23.10) [27.20-36.19]	11.4	0.58	✓
Food on Floor	55.84 (29.51) [49.83-61.85]	57.07 (32.88) [50.34-63.81]	1.23	0.04	x
Ceiling	32.01 (30.82) [26.10-37.92]	81.41(21.39) [77.25-85.57]	49.40	1.64	✓
Wall hazards (not in open view)	11.59 (15.56) [8.61-14.57]	56.70 (27.86) [51.28-62.12]	45.11	2.90	✓
Shelf	39.71(15.67) [36.71-42.71]	45.56 (18.35) [41.99-49.13]	5.85	0.37	x
Mouse Traps	10.51 (21.73) [6.35-14.68]	31.01 (30.67) [25.05-36.97]	20.50	0.94	✓
Pictogram	47.77 (30.26) [41.80-53.75]	55.18 (32.11) [48.72-61.66]	7.42	0.25	x
Ceiling Tiles Moved	40.89 (45.47) [32.17-49.60]	87.98 (28.93) [82.35-93.61]	47.09	1.04	✓
Smoke Detector Covered	35.98 (47.48) [26.88-45.08]	65.38 (46.77) [56.29-74.48]	29.40	0.62	✓

Large Sized (more than .0.5m)	40.31 (28.75) [34.80-45.82]	50.73 (20.93) [46.66-54.81]	10.43	0.36	x
Small Sized (less than.0.5m)	38.48 (12.06) [36.16-40.79]	53.10 (14.17) [50.36-55.84]	14.62	1.21	✓
Bare Wires	27.72 (44.99) [18.84-36.60]	65.98 (47.62) [56.38-75.58]	38.26	0.85	✓
N Hazards Written	31.22 (10.36) [29.24-33.21]	48.86 (12.92) [46.34-51.37]	17.64	1.70	✓

**** using an Independent t-test with a Bonferroni correction. ✓= was significant x= was not significant***

5.1.3 The Distribution of Visual Inspection Scores

The histograms in Figures 11 & 12 below, show the control and experimental group distributions for VIP as measured by the mean percentage hazards observed. As such it is the first time a percentage density function has been presented for the observation of workplace hazards. What can be seen from these histograms is that they were both platykurtic distributions in shape (a negative kurtosis or sharpness of peak, with light tails). In contrast the control group showed a slight positive skew (the mean and tail is further towards the left on the x axis) which became a more positive skew for the experimental group.

Figure 11 Control Group; Visual Inspection Performance Distribution

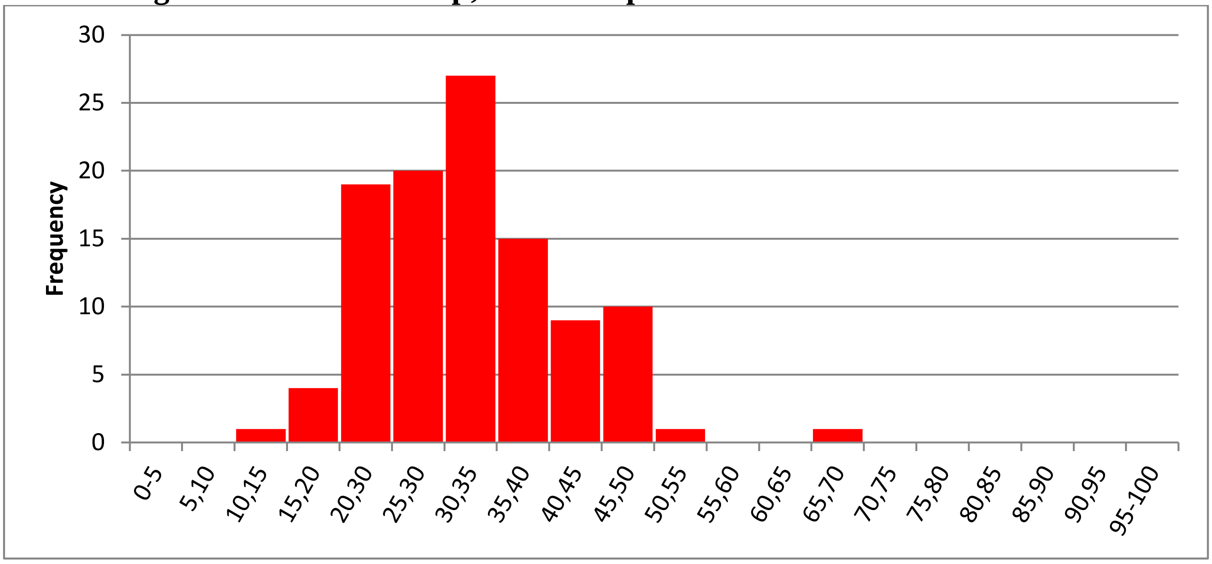
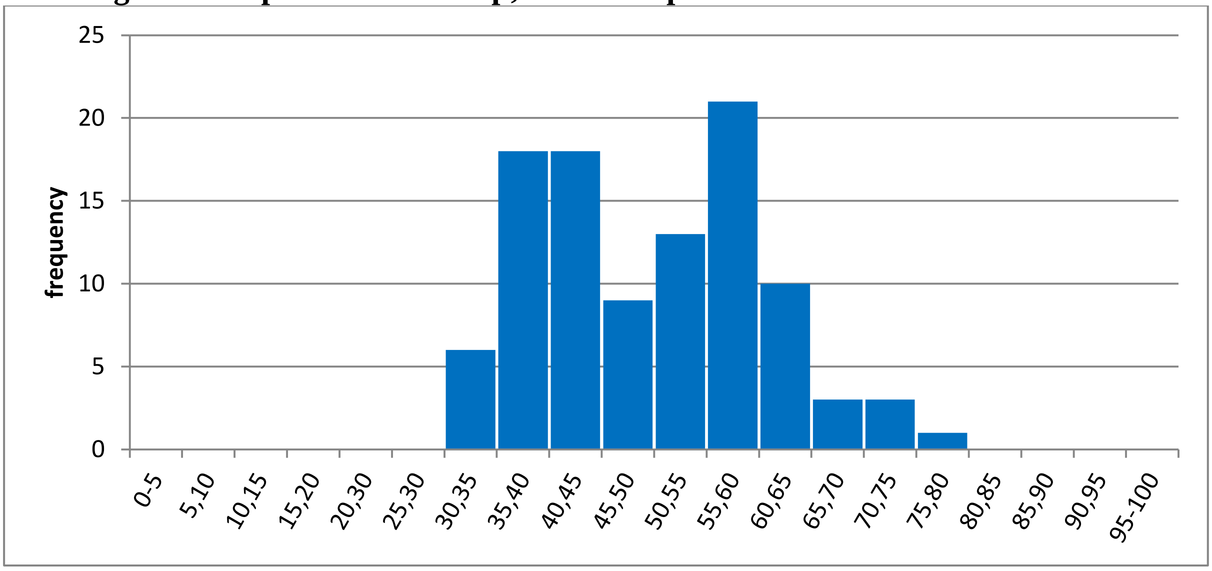


Figure 12 Experimental Group; Visual Inspection Performance Distribution



A Shapiro Wilks test was conducted and normality was not confirmed for either the control, or experimental group results. However as Field, (2013) points out, the large sample size of each treatment group still allows inferential statistical tests to be conducted. Furthermore the SPSS derived Q-Q plots demonstrated how a handful of scores “prevented” normal distribution for both groups. A case can be made for removing these outliers, but it would mean a loss of data. Furthermore, as normality can be assumed given the large sample size (Field, 2013) there does not seem to be any advantage from removing any data.

It is also noted that the control group results show a slightly positively skewed distribution, but that experimental group results showed a less positive skew. When described by the skewness statistic calculated by SPSS, the results can be seen in Table 16 below.

Table 14 Measure of Skewness; Visual Inspection Performance Distribution

Treatment Group	**Skewness Statistic**	**Standard Error**
Control	0.69	0.23
Experimental	0.22	0.24

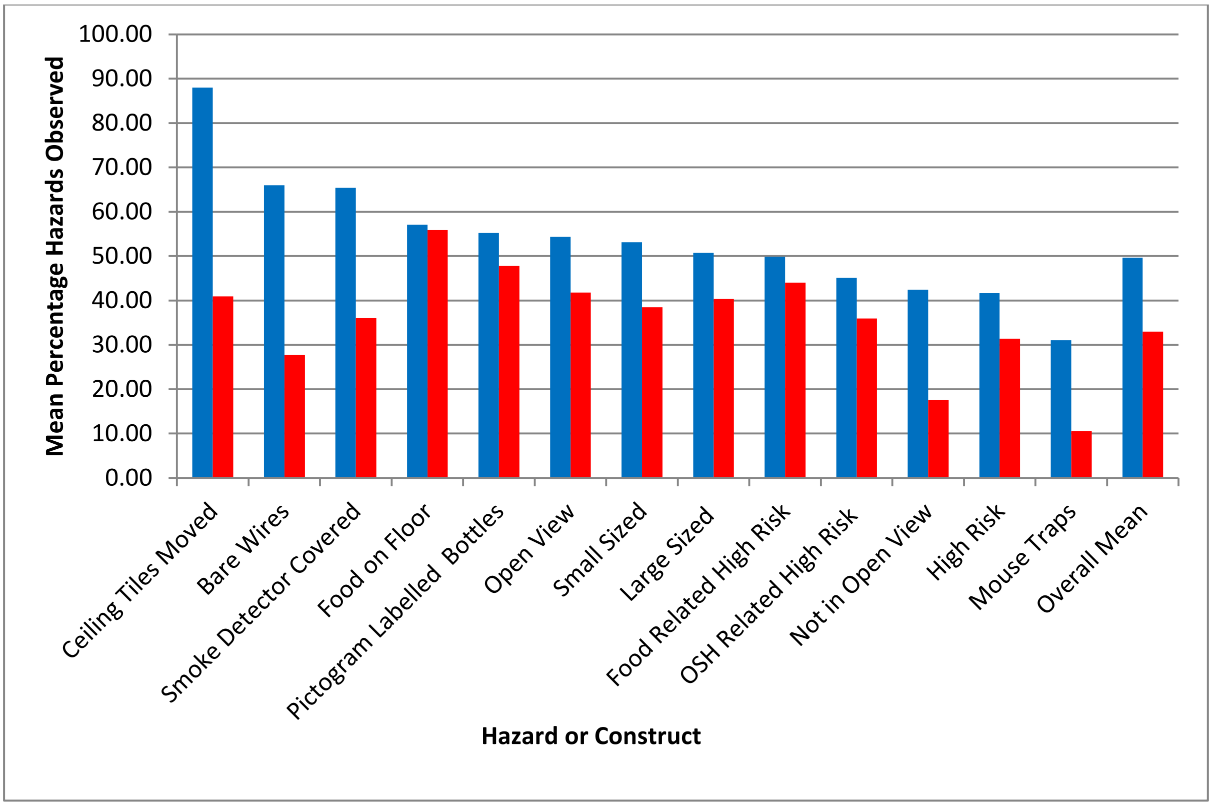
In displaying the percentage density functions for VIP, Figures 11 & 12 again demonstrate the strength of the intervention as well as graphically describing the platykurtic distribution shape. These distribution shapes have not been reported in the literature and could be describing human visual inspection variability.

5.1.4 Visual Inspection Performance by Hazards & Constructs

The first analysis conducted on the overall data was to break down the hazards into specific hazards or constructs. This allowed any differences between control and experimental participants to be clearly seen. Although all this data is all available from Table 15 above, a bar chart representation of selected hazards and constructs allows a visual appreciation of the main findings. How many of the 24 hazards and constructs in Table 15 were chosen for presentation will now be detailed. The first consideration was the number of hazards or constructs to display. Appendix 6 lists the total number of 27 hazards that could potentially have been present and thereby observed within the trials (with some hazards being present more than once). The mean number of hazards per trial was 36.17 with a SD of 4.04. There was also an additional dependent variable created by the methodology itself, that of the number of hazards written down by participants.

So the first step in considering which hazard and possible constructs to present from the many generated, was to review them all as possible contenders for presentation, showing the effects from the intervention. This review demonstrated that were some hazards and constructs that stood out in terms of illustrating the findings and effects of the intervention when considering the aim, research hypothesis and question. Therefore, those hazards and constructs with the highest and lowest levels of VIP illustrated the range of effects from of the intervention and were included. By way of contrast, those hazards and constructs that exhibited the least amount of influence from the SVS method were included to highlight the limits of the intervention. The construct; hazards in open view, which again displayed a relatively small effect from the intervention was selected. As a comparison, hazards not in open view was also selected. Similarly, small sized hazards (less than 0.50m in any direction) was selected, prompting the construct large sized hazards (more than 0.50m in any direction) to be picked. Some measure on the intervention's effect on high risk hazards was also deemed important in terms of generalising the potential impact of the findings to the wider EHS community. Three constructs; high food risk hazards, high OSH risk hazards and an overall high risk construct, were thereby selected for presentation. The following bar chart (Figure 13) presents this graphical display showing the effect of SVS on selected hazards and constructs. For noting and as used throughout this thesis; control group data are shown in red and experimental group data are shown in blue.

Figure 13 Visual Inspection Performance for Selected Hazards & Constructs



As Figure 13 uses the data from Table 15 this bar chart again paints a very clear picture of the beneficial effect of using SVS. Figure 15 serves a useful purpose here, that of graphically displaying the beneficial effect of SVS use when compared to the results from control group participants.

5.1.5 What Predicted Successful Observation?

The next stage of the analysis, was to see if there were any salience characteristics associated with any hazards that may have predicted observation. Some hazards were close in terms of VIP for example, food on the floor was observed by 55.84% of control group participants and 57.07% of experimental participants. In effect, this was roughly the same VIP rate, as there was no statistically significant differences between treatment groups. However, and after much reflection, there were no salience characteristics found. This will be further discussed in section 6.4.1 In contrast, and as can be seen in Table 15 and graphically shown in Figure 14 below, whether or not the hazard was in open view, did have a significant effect.

Figure 14 VIP for Open View & Obfuscated Hazards

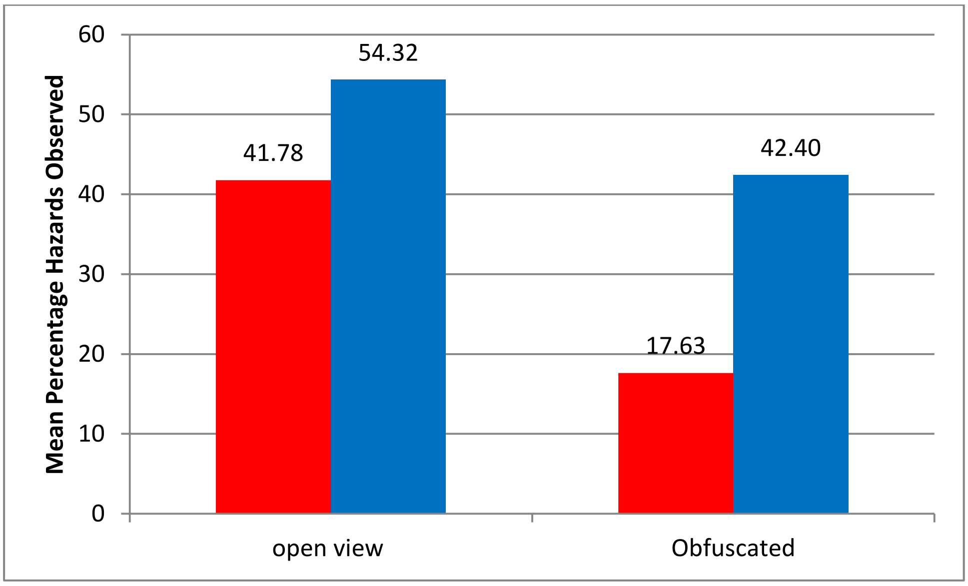


Figure 14 suggests that one effect from SVS use is to promote visual search in those areas that require effort to fully observe. Looking behind, around and underneath items was emphasised during the SVS training session and the results are clear to see in Figure 14. However the nature of the intervention needs to be born in mind when considering this finding. It does not suggest that control group participants can’t see obfuscated hazards, but it does suggest they don’t go looking up, behind, under and around items as much as their experimental group colleagues. In short anywhere that required effort to look, tended to be ignored more by control group participants. In summary experimental group participants using the SVS method observed more hazards by; matching or slightly bettering their control group colleagues with open view hazards. But they substantially and significantly exceeded their control participants by; looking up, behind around and underneath items in the kitchen, and thereby observed more hazards in those areas that needed a bit more effort to see clearly.

5.1.6 The Number of Hazards Written Down

What can also be seen in Figures 15 & 16 below is the beneficial effect of SVS on the number of hazards written down. This can be seen with the platykurtic distributions for both the control and experimental groups below.

Figure 15 Distribution of N Hazards Written Down by the Control Group

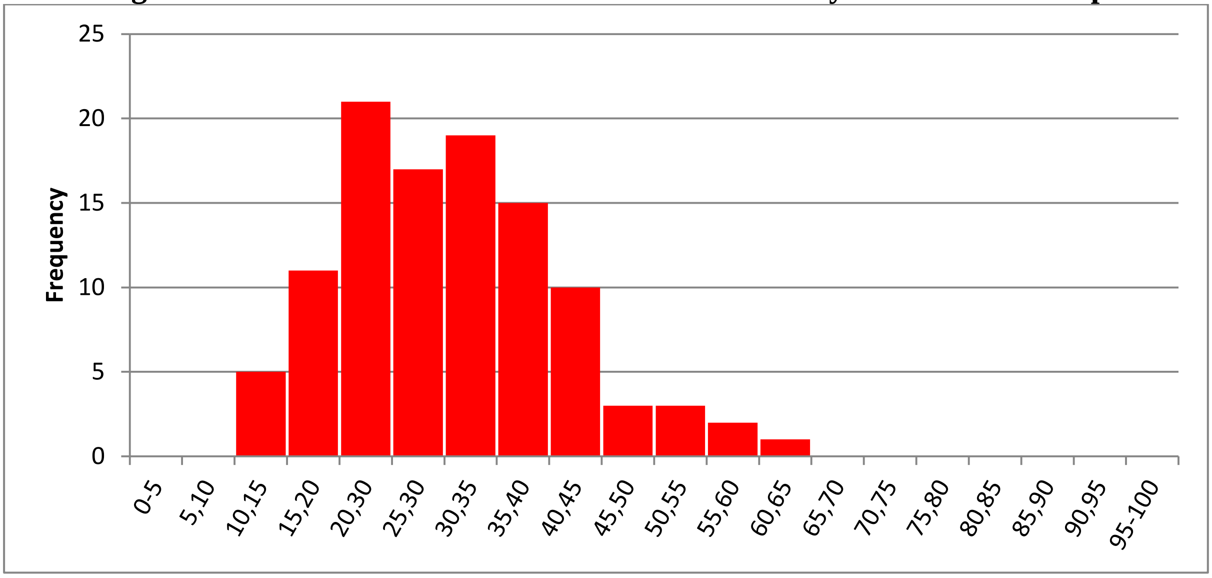
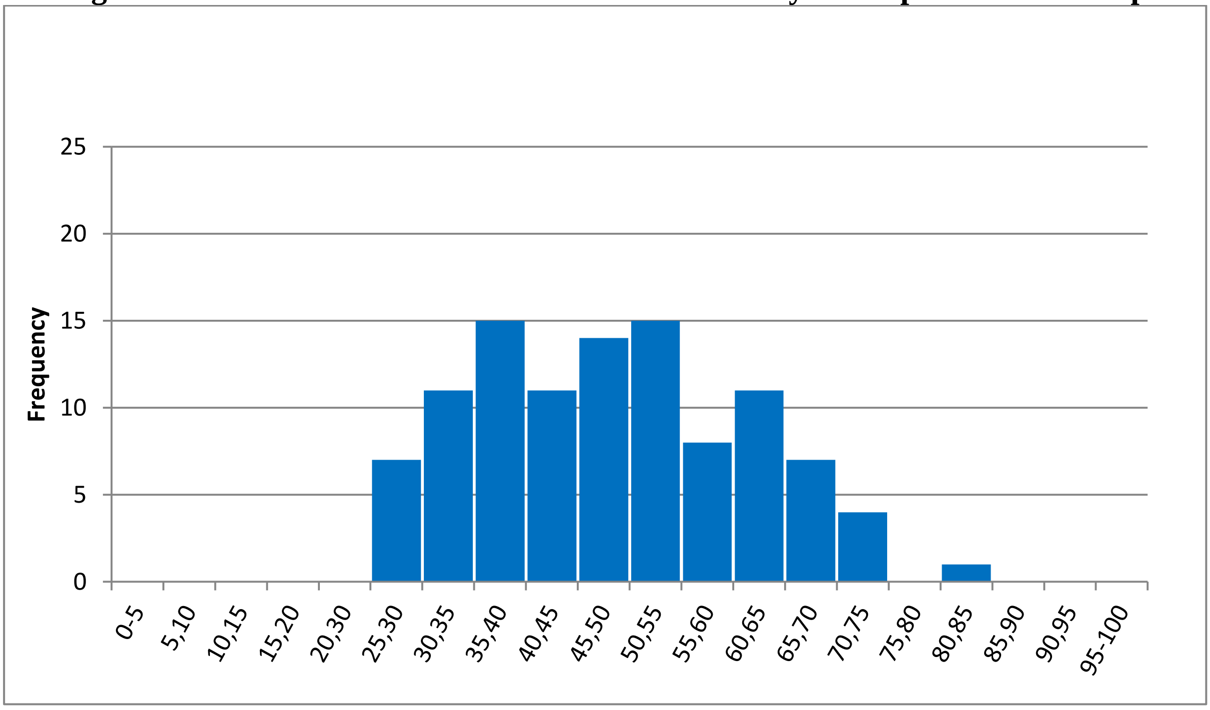


Figure 16 Distribution of N Hazards Written Down by the Experimental Group



A Shapiro Wilks test shows that these two distributions, returned *p* values of .01 and .09 for control and experimental groups respectively. This shows normality for the experimental group but not for the control group. But as mentioned previously, the large sample size can allow normality to be assumed (Field, 2013). The skew results are demonstrated in Table 17 below.

Table 15 Measure of Skewness of Number of Hazards Written Down

Treatment Group	**Skewness Statistic**	**Standard Error**
Control	0.62	0.23
Experimental	0.28	0.24

As can be seen, experimental group participants wrote down a mean 17.64 more hazards than their control group colleagues (see Table 15). This increase was significant, and once again came with a large effect size. These results also demonstrated an efficiency gain by SVS users over their control group colleagues. SVS users observed more hazards, but in writing more hazards down, consequently had less inspection time than their control group colleagues. This effect has not been noted previously in the visual search literature.

5.2 The Kitchen Study; Further Inferential Results

The strength of SVS intervention has already been described by the data from Table 15 to include effect sizes. A further measure of the intervention's effect will now be presented in the form of regression calculations.

5.2.1 Regression Analysis for Visual Inspection Performance

There was a simple linear regression calculation conducted by SPSS, whereby the dependent variable was VIP for each treatment group. This was done to give an additional measure of the intervention's effect. As Field, (2013) explains; a simple linear regression will give rise to a line of best fit between two variables. The equation of a straight line; $y = mx + c$ can be used to model the line of best fit between two variables. In this form, y is the dependent variable, m is the gradient of the line, x is the independent variable and c is the intercept on the Y axis.

As Field, (2013) explains, the equation of a straight line can also take the form; $y = (b_0 + b_1X)$ and these beta values are outputted by SPSS. Using this form of the equation then $y =$ the dependent variable (visual inspection performance), X is the independent variable (if systematic visual search was used), b_1 is the slope of the gradient and b_0 is the intercept (Field, 2013). Table 18 below, details these regression results in a format recommended by Field, (2013).

Table 16 Regression Results for the Intervention

	b_0	SE b_0	b_1	p	R^2
Constant	32.97	0.97		≤.001	
BCa 95% Confidence Interval	[31.36-34.59]				
Participant Treatment group	16.67	1.41	0.64	≤.001	0.41
BCa 95% Confidence Interval	[13.96-19.63]				

BCa; Bias corrected and accelerated by SPSS v 21

The high value of R^2 again demonstrates the strength of SVS in terms of improving visual inspection performance. Furthermore as $R^2 = 0.41$; then 41% of the variance in VIP is due to using the SVS method.

5.2.2 Regression Result for the Number of Hazards Written Down

The following Table 19 shows this strong effect again, this time for the number of hazards written down.

Table 17 Regression Results for the Intervention on N Hazards Written

	b_0	SE b_0	b_1	p	R^2
Constant	31.22	1.61		≤.001	
BCa 95% Confidence Interval	[29.25 – 33.07]				
Participant Treatment group	17.63	1.13	0.60	≤.001	0.37
BCa 95% Confidence Interval	[14.24 - 21.32]				

BCa; Bias corrected and accelerated by SPSS v 21

5.2.3 The Effects from Non Controlled Extraneous Variables

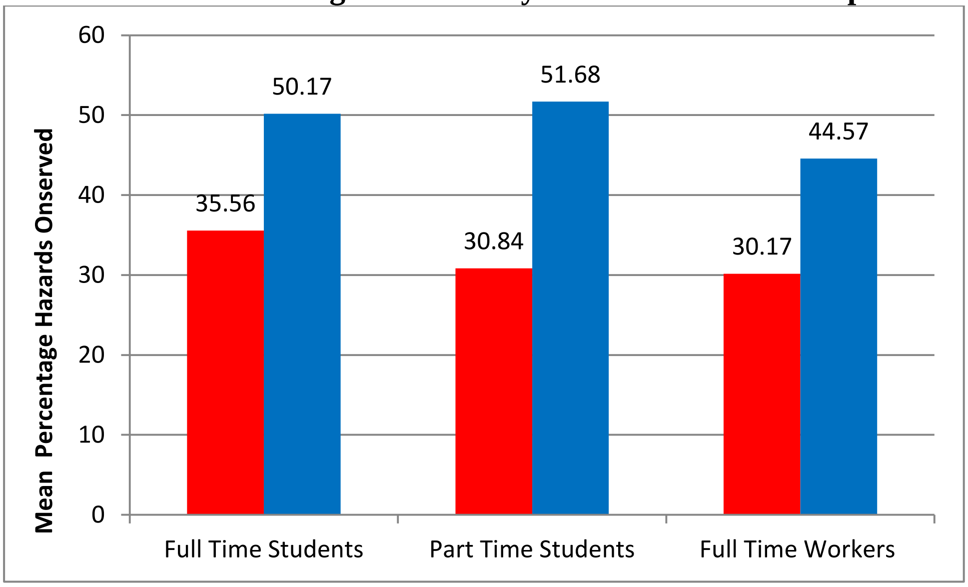
As mentioned, the results in Table 15 together with the robust experimental design give weight to evidence for SVS being beneficial when compared to customary visual inspection practice. However, the results could have been biased by non-controlled extraneous variables such as the kitchen used or participant experience. So the next stage of the analysis was to see if there was any effect on VIP from extraneous variables. In this way, additional confidence that the SVS intervention was responsible for the effects on VIP could be reported.

This analysis was conducted by splitting the 211 participant sample into their treatment groups and applying statistical treatments to each group separately. The effect of this manipulation, as recommended by Tabachnick & Fidel, (2013), is to allow further analysis on the separate samples without having to control for the intervention effect. To recap, for the kitchen study there were seven extraneous variables. The mathematical treatment to evaluate the effect of these extraneous variables on VIP was ANOVA and χ^2. This approach is recommended by Field, (2013), who also advises that graphical representations of such variables, are a useful first step in any analysis.

5.2.4 Visual Inspection Performance by Work status of the Participant

The effect on VIP from the work status of participants can be seen in Figure 17 below.

Figure 17 VIP by Work status of Participant



In Figure 17, the participants by treatment group (red bars; control, blue bars; experimental) were further categorised into full time student, part time student and full time worker status in order to compare their respective visual inspection performances. To see if there were any significant differences in VIP between these categories, a one way ANOVA was conducted in SPSS. This “post hoc” multiple comparison method used the Tukey and Games-Howell tests, as recommended by Field, (2013).

The ANOVA calculation found there was a significant 5.39% difference in VIP between control group full time students, and their full time professional colleagues. The ANOVA statistics to demonstrate this were;

$F = 4.28, (2,104)\ p = .016\ \ \eta = .28$

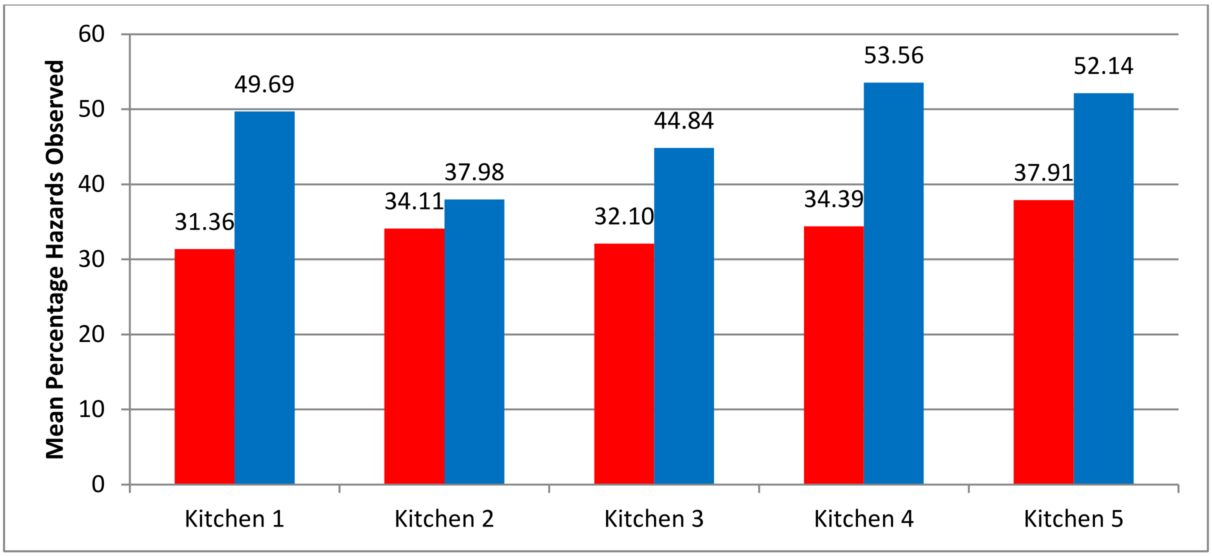
This finding represents a low to medium effect size. Field, (2013) states that; η is the same statistic as r, and so measured from -1 to +1. Furthermore, the difference in VIP between part time students and professional participants was not significant. Although not causal due to the experimental design not independently manipulating this categorical extraneous variable, it seems intuitive that the type of student predicts VIP, and not the other way round. So the question is why would control group full time students, have a 5% or so higher VIP, than their control group working colleagues who have far more experience in terms of years worked. Two reasons could account for this counter-intuitive finding. Firstly, control group full time students could be exhibiting a beginner's naivety in over rating the importance of observed hazards. Put another way, they could be exhibiting an over zealousness in seeing hazards where they don't exist. Secondly, it could be that their control group working colleagues are exhibiting outcome or confirmation bias (Kahneman, 2011) whereby they do not regard the hazards they observed as actual hazards, or they only see what they want to see, a superficially clean kitchen. This point is further discussed in section 6.5.2.

In contrast, experimental group participants had differences in VIP between the categories of participants, but they were not significant. All of these differences are around five percentage points which in the overall scheme of things, is not a great deal considering the effect of internal bias (see section 6.2). However the differences were statistically significant for the control group full time students when compared to control group full time working participants.

5.2.5 Effect of Kitchen Used on Visual Inspection Performance

The effect on VIP from the kitchen inspected by participants, can be seen in Figure 18 below.

Figure 18 Effect of Kitchen Used on VIP



In Figure 18 above, the effect of the kitchen on participant VIP was analysed (red bars; control, blue bars; experimental). This was to see if there were any significant differences in VIP due to the kitchen being inspected. Again a one way ANOVA was conducted in SPSS. This "post hoc" multiple comparison method used the Tukey and Games-Howell tests, as recommended by Field, (2013).

The differences in VIP for control group participants was not significant and so there was no effect found from the kitchen used. However, the kitchen used was found to have a significant effect on experimental group participants. Here, kitchen 2 was found to have a significantly different VIP amongst participants when compared to the resulting VIP from the four other kitchens. The ANOVA statistics to demonstrate this effect were;

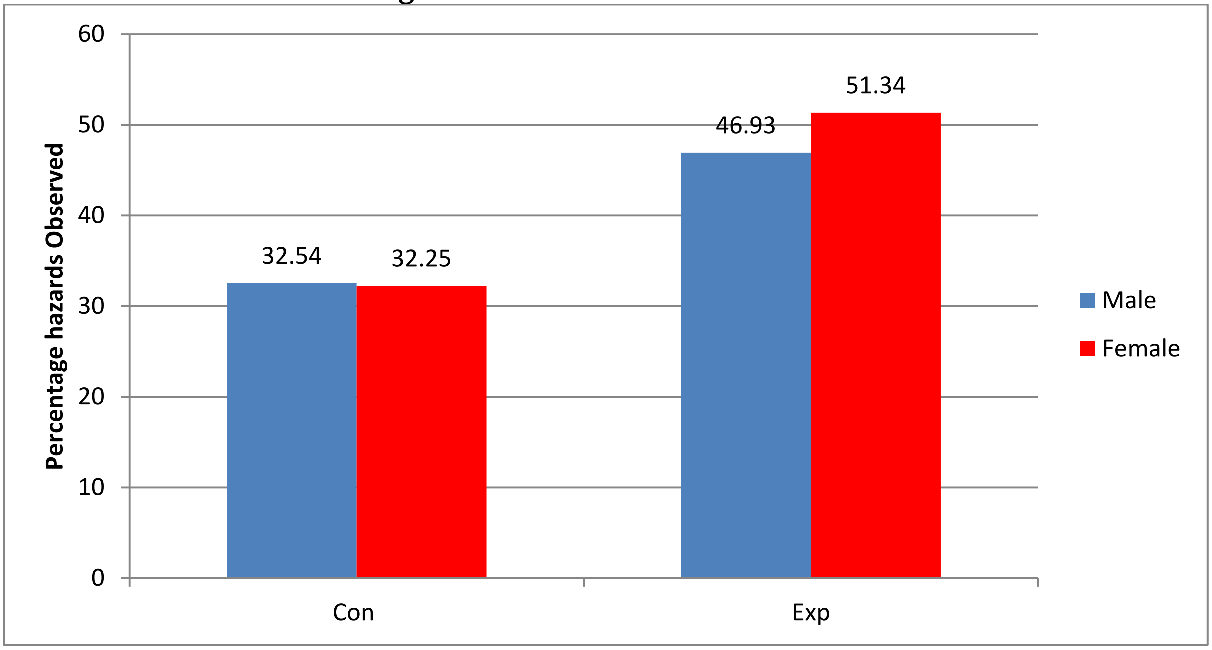
$F\ (4, 99) = 3.58,\ p= .009\ \ \eta=.36$

Kitchen 2 was found to have a significant medium to strong effect on VIP. Experimental group participants using this particular kitchen returned a VIP of 37.98%. The average VIP for the remaining four kitchens was 50.05% This 12.07% difference could be due to some characteristic in the way this kitchen was set up, or simply sample error. Of the two explanations, sample error is more likely. Kitchen 2 was only used in one trial, and by six experimental participants (see Table 12). Furthermore, if kitchen 2 was systematically affecting VIP, a similar effect would have been seen with control group participants who would have also demonstrated a significantly lower VIP. As this was not the case, sample error seems to be the most likely explanation.

5.2.6 Effect of Gender on Visual Inspection Performance

The effect on VIP from participant gender can be seen in Figure 19 below.

Figure 19 Effect of Gender on VIP

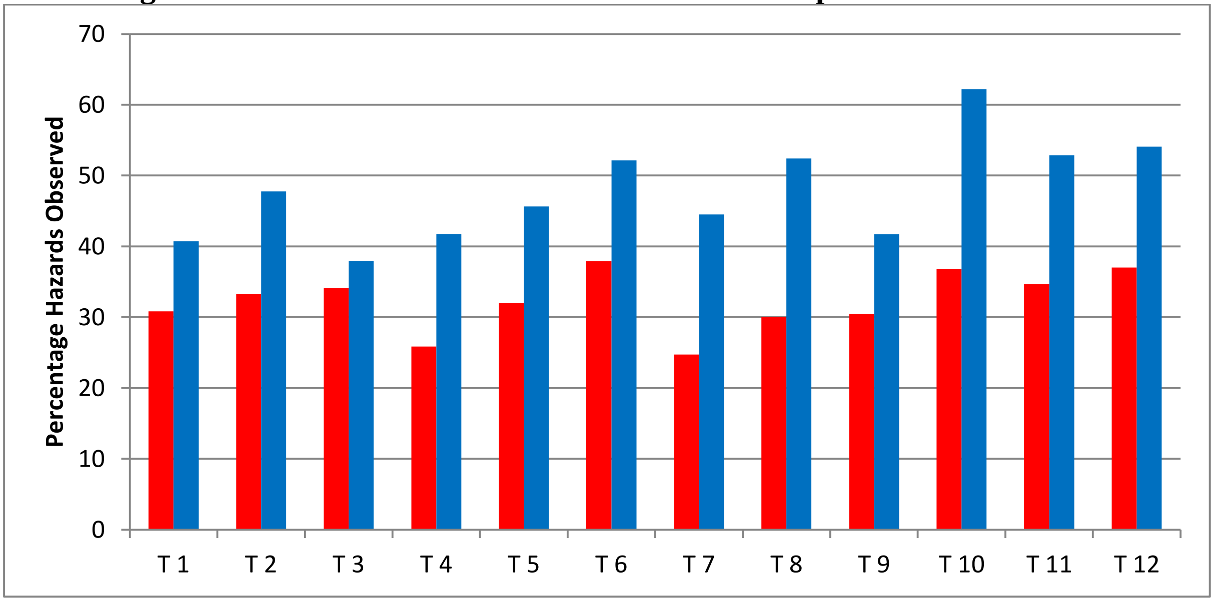


VIP by female experimental group participants, was 4.41% better that their male colleagues. However this finding was not significant when using a χ^2 test. Therefore there is no evidence to suggest that gender had a significant effect on VIP by control or experimental group participants.

5.2.7 The Effect of Trial Number on Visual Inspection Performance

The effect on VIP from the trial the participants took part in can be seen in Figure 20 below.

Figure 20 The Effect of Trial Number on Visual Inspection Performance



In Figure 20 above, the effect of the particular trial undertaken by participants and their resultant VIP was analysed (red bars; control, blue bars; experimental).. This was to see if there were any significant differences in VIP due to the trials themselves. Again a one way ANOVA was conducted in SPSS. This “post hoc” multiple comparison method used the Tukey and Games-Howell tests, as recommended by Field, (2013).

The trial number did systematically affect VIP in both treatment groups with statistical significance. For the control group, the strength effect was low to medium ($\eta = .46$);

$F\ (11, 95) = 2.36, p = .013,\ \eta = .46$

For the experimental group the trial number strength of effect was large ($\eta = .62$);

$F\ (11, 92) = 5.33, p = .001,\ \eta = .62$

What these results are saying is that different trials had participants that varied in their visual inspection performances. As participants were usually from the same class or group, it reflects that some cohorts were better or worse than others in terms of VIP. However it should not come as any surprise that different trials returned varying mean VIP, as different individual participants also returned varying VIP results. In summary it was demonstrated that there were no experimental design issues that were systematically affecting both treatment groups. There were some effects from the kitchen used and work status, but again they did not affect both treatment groups at the same time. Having found that in overall terms, internal bias from the experimental design was limited in affecting VIP the next stage was to consider correlations between relevant variables.

5.2.8 Kitchen Study Correlations

Correlations measure the strength of any relationship between two variables (Field, 2013). Whilst any two interval ratio variables can be correlated, the variables cannot be co-dependent. So it is normal practice to correlate using extraneous or independent variables with dependent variables. There is also the proviso here, that they both need interval ratio data although some scholars use ordinal variables (Field, 2013). Therefore the correlational analysis in Table 20 below, is evaluating possible relationships between VIP and extraneous variables of; N Hazards Written, N Hazards Present, N Years Working and N Hazards/m^3.

Table 18 Correlations of Visual Inspection Performance With Extraneous Variables

	N Hazards Written	**N Hazards present**	**N Years Working**	**N Hazards/m^3**
Control Group VIP	0.29** CI [0.17-0.45] *P*=.002	-0.048 CI [-0.28-0.13] *P*=.62	-0.18 CI[-0.31--0.06] *P*=.06	0.01 CI [-0.18-0.22] *P*=.89
Experimental Group VIP	0.52** CI [0.39-0.64] *p* = ≤.01	≤.001 CI[-0.19-0.16] *P*=.10	-0.21* CI[-0.35--0.04] *P*=.03	-0.12 CI[-0.33-0.09] *P*=.22

* *p*= ≤.05 ** *p* = ≤.01

The correlation between VIP and the number of hazards written down for both treatment groups was significant. This suggests that for all participants, the better the VIP of the participant, the higher the number of hazards they wrote down. Whilst this would seem an intuitive result, there was a difference in the strength of this relationship by treatment group. For the control group, the strength of relationship between VIP and the number of hazards written down was small, but was large for experimental participants.

The other finding seen in both treatment groups was the negatively correlated relationship between VIP and the number of years worked. This result was unexpected and counter intuitive as this statistic is saying that; the less work experience the participant had, the better his VIP. Whilst this correlation mirrors Figure 17, (VIP by work status), this result needs to be treated with some caution as significance was only returned with experimental group participants. Nevertheless a negatively inverted small to medium effect was seen. As mentioned earlier, this finding could be explained by over zealous students simply writing down more hazards which by chance were correct, or it could be the effect of confirmation or outcome bias amongst control participants.

5.2.9 Summary of Main Findings from the Kitchen Study

It was found that SVS significantly increased VIP from a mean of 32.97% to 49.54%. This VIP improvement, was found to have a large effect size (Cohen’s *d* = 1.85). This improvement was seen across all hazards and constructs as demonstrated in Table 15 and seen in Figure 13. The least amount of VIP improvement occurred with hazards that were in open view. This 16.68% improvement is underscored when considering it was achieved with only thirty minutes training. A similar beneficial effect was also seen in in terms of the number of hazards written down. SVS users wrote down 17.64% more hazards than their control group colleagues. This also represented a large effect size (Cohen’s *d* = 1.70).

For control group participants, whether or not the observer had a clear line of sight to the hazard, was the only (and in hindsight) rather obvious predictor of visual inspection performance. Once the effect of SVS is stripped out, there does not seem to be any single clear salience characteristic of the hazards themselves that makes them any more likely to be observed by either control or experimental participants. This leads to the suggestion that SVS works by utilising a behavioural visual search algorithm that promotes an exhaustive observation of the area under analysis. This issue will be further discussed in section 6.4.

5.3 To Reject or Not Reject the Null Research Hypothesis

At this point, having presented the descriptive and inferential statistics for the kitchen study it is time to address an important scientific aspect of this thesis; which research hypothesis to reject or not reject. To recap from section 4.5.4, the null research hypothesis (H_0) is; the use of SVS will ***not*** result in different visual inspection performance rates when compared to current practice. The alternative hypothesis (H_1) is; the use of SVS ***will*** result in different visual inspection performance rates when compared to current practice.

Phrased in these terms; H_0, can be represented by the VIP rate for control group participants in the kitchen study at 32.97%. The SVS intervention had the effect of increasing VIP by a mean 16.68% ($p \leq .001$).

Therefore, the results demonstrate strong evidence to state that for the kitchen study; SVS is beneficial for current practice, when looking for observable workplace hazards during visual inspections. Therefore H_0, can be rejected in favour of the alternative hypothesis H_1.

Furthermore, the effect of extraneous variables were evaluated by ANOVA with little effect being found. In addition, the robust experimental design used the “gold standard” randomised controlled trial methodology, demonstrated comparability of control and experimental participants, ensured a high degree of ecological validity and achieved the requisite statistical power of .08. In summary, there can be statistical confidence in support of H_1.

5.4 The Aviation Study; Descriptive Results

This section will present the detailed data generated by the experimental design from the aviation study. As with the kitchen study, it will begin by presenting descriptive data and statistical tests before proceeding to inferential data and testing.

5.4.1 A Demographic Comparison Between Treatment Groups

As mentioned in the Chapter 4 (Methodology) the aviation study was smaller. The number of participants in the aviation study was considerable lower at 26 compared to the kitchen study at 211. Therefore the extent of the demographic data will be correspondingly less. Table 21 below, presents these comparative data. It should be noted that all the participants were male. As per the kitchen study, the number of years working was also recorded.

Table 19 Demographic Comparison of Treatment Groups in the Aviation Study

	Control	**Experimental**
N full time professionals (all males)	13	13
Years working in the job	Mean= 24.08 SD= 6.36	Mean = 25.08 SD = 4.86

What can be seen here is the closeness in demographic characteristics between control and experimental groups in terms of gender, job status and years of experience. They were similar enough in all categories to warrant a valid comparison and ensure any effects on the 13 experimental participants, could be attributed as far as possible, to the intervention. As expected a χ^2 test was found no significant differences between mean work experience of treatment group participants ($p = \leq .05$)

5.4.2 Descriptive Results from the Aviation Study

Appendix 6 lists the 29 hazards and plants on, and within, the light aircraft together with a description of their importance categorised as high or low. The type of hazards on the light aircraft were typical aviation defects such as; loose screws, split pins left unclamped, locking wires cut and areas of corrosion. “Plants” included screws partially unscrewed, circuit breakers left open and loose items left unsecured such as a mobile phone, coins, pens and a screwdriver. These loose items can interfere with pilots during aerial manoeuvres and are deemed high risk. Table 22 below, presents the overall results with Cohen’s *d* and significance testing. It should be noted that the control and experimental columns present the mean percentage hazards observed (VIP) for each category of hazard listed in the first column.

What Table 22 below demonstrates, is SVS’s beneficial effect when compared to the control group in terms of overall VIP demonstrated. The comparability of participants, and the robust experimental design, also strongly supports the theory that the beneficial effect on VIP, was caused by the use of SVS. The difference in overall mean number of hazards observed by control and experimental groups was again significant at 25.73% and with a large effect size (*p*=≤.0016). In contrast to the kitchen study, some hazards and plants were observed at a higher rate by control group participants, a point that will be detailed in section 5.4.4 below. The highest level of VIP was the “plant” a flat tyre, was seen by all participants. The lowest level of VIP occurred with plant 7, the starboard wing fuel cap earth wire removed. With this plant, only one participant in each treatment group (M=7.69% SD = 27.74) observed the particular defect.

Table 20 VIP Data by Specific Hazards & Constructs; Aviation Study

Hazards Identified	Control (VIP M&SD) [95% CI]	Experimental (VIP M&SD) [95% CI]	%age difference in M’s	Cohen’s *d*	p= ≤.0016*
Overall Mean	37.93 (9.65) [32.75-43.10]	63.66 (10.11) [58.39-69.42]	25.73	2.67	✓
Hazard 1	23.07 (43.85) [-3.42-49.57]	61.53 (50.63) [30.93-92.13]	38.46	0.88	x
Hazard 2	38.46 (50.63) [7.86-69.06]	61.53 (50.63) [30.94-92.14]	23.08	0.46	x
Plant 1	46.15 (51.89) [14.80-77.51]	76.92 (43.85) [50.42-103.42]	30.77	0.59	x
Plant 2	76.92 (43.85) [50.42-103.42	58.85 (51.89) [22.49-85.20]	-23.08	-0.53	x
Plant 3	69.23 (48.04) [40.20-98.26]	61.54 (56.64) [30.94-92.13]	-7.69	-0.16	x
Plant 4	15.38 (37.55) [-7.3-38.08]	61.53 (50.64) [30.94-92.14]	46.15	1.23	✓
Hazard 3	61.53 (50.64) [30.94-92.14]	92.31 (27.34) [75.55-109.07]	30.77	0.61	x
Plant 5	53.85 (51.89) [22.49-85.20]	53.84 (51.89) [22.49-85.20]	0	0	x
Plant 6	30.76 (48.04) [1.74-59.80]	38.46 (50.64) [7.86-69.06]	7.69	0.61	x
Hazard 4	38.46 (50.64) [7.86-69.06]	53.85 (51.89) [22.49-85.20]	15.38	0.30	x
Plant 7	7.69 (27.74) [-9.07-24.45]	7.69 (27.74) [-9.07-24.45]	0	0	x
Hazard 5	46.15 (51.89) [14.80-77.51]	46.15 (51.89) [14.80-77.51]	0	0	x
Plant 8	53.84 (51.88) [22.49-85.20	69.23 (48.04) [40.20-98.26]	15.38	0.30	x
Hazard 6	7.69 (27.74)	76.92 (43.85)	69.23	2.50	✓

	[-9.07-24.45]	[50.42-103.42]			
Hazard 7	15.38 (37.55) [-7.31-38.08]	69.23 (48.04) [40.20-98.26]	53.85	1.43	✓
Plant 9	69.23 (48.04) [40.20-98.26]	84.61 (37.55) [64.92-107.30]	15.38	0.32	x
Plant 10	53.85 (51.88) [22.49-85.20]	46.15 (51.89) [14.80-77.51]	-7.69	-0.15	x
Hazard 8	7.69 (27.74) [-9.07-24.45]	69.23 (48.04) [40.20-98.26]	61.54	2.22	✓
Hazard 9	15.38 (37.55) [-7.31-38.08]	53.85 (51.89) [22.49-85.20]	38.46	1.02	x
Hazard 10	23.08 (43.85) [-3.4-49.58]	30.77 (48.04) [1.74-59.80]	7.69	0.18	x
Hazard 11	100 (0) [100-100]	100 (0) [100-100]	0	0	x
Hazard 12	15.38 (37.55) [-7.31-38.08]	7.69 (27.74) [-9.07-24.45]	-7.69	-0.20	x
Plant 11	69.23 (48.04) [40.20-98.26]	84.62 (37.55) [61.92-107.31]	15.38	0.32	x
Plant 12	7.70 (27.74) [-9.07-24.52]	84.62 (37.55) [61.92-107.30]	76.92	2.77	✓
Hazard 13	38.46 (50.64) [7.86-69.06]	76.92 (43.85) [50.42-103.42]	38.46	0.76	x
Plant 13	76.92 (43.85) [40.42-103.42]	76.92 (43.85) [40.42-103.42]	0	0	x
Plant 14	23.07 (43.85) [-3.42-49.58]	92.30 (27.74) [75.55-109.07]	69.23	1.58	✓
Plant 15	7.69 (27.73) [-9.07-24.45]	84.62 (37.55) [61.92-107.31]	76.92	2.77	x
Hazard 14	23.08 (43.85) [-3.42-49.58]	69.23 (48.04) [40.20-98.26]	69.23	1.05	x
%age of important hazards observed	46.15 (18.36) [35.06-57.25]	68.53 (16.13) [59.39-77.68]	22.38	1.22	x
N Hazards Written	26.46 (9.82) [20.52-32.40]	52.15 (12.43) [44.64-59.66]	25.66	2.62	✓

**** using an Independent t-test with a Bonferroni correction. ✓= significant x = not significant***

5.4.3 The Distribution of Visual Inspection Scores

The histograms in Figures 21 & 22 below show the control and experimental group distributions for VIP.

Figure 21 Control Group VIP Distribution

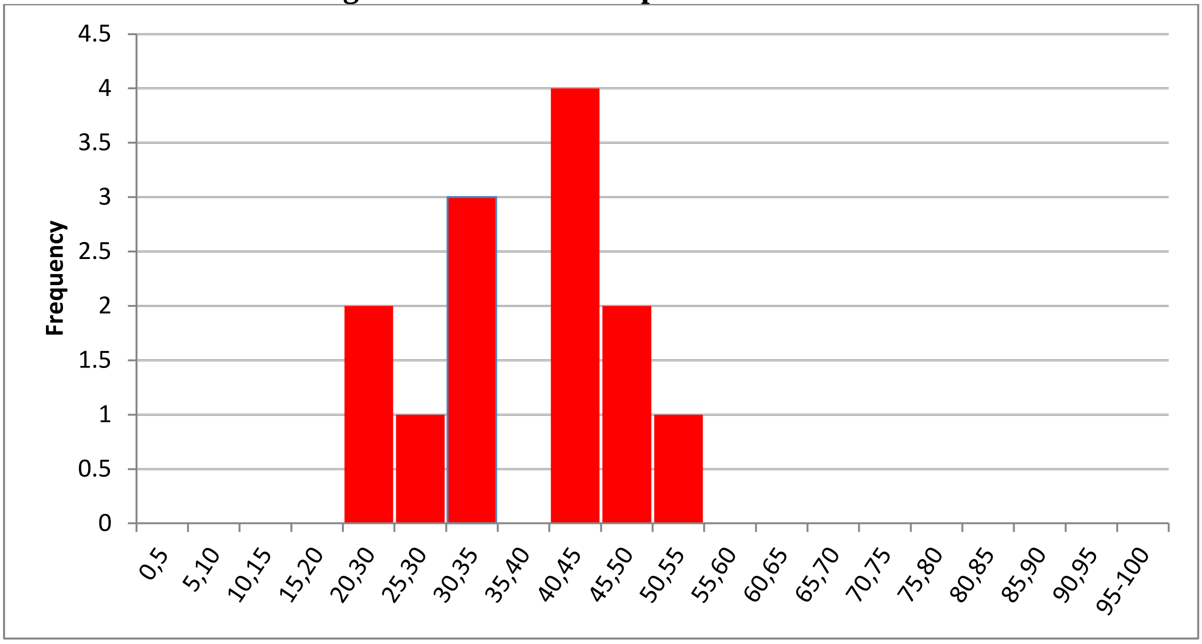
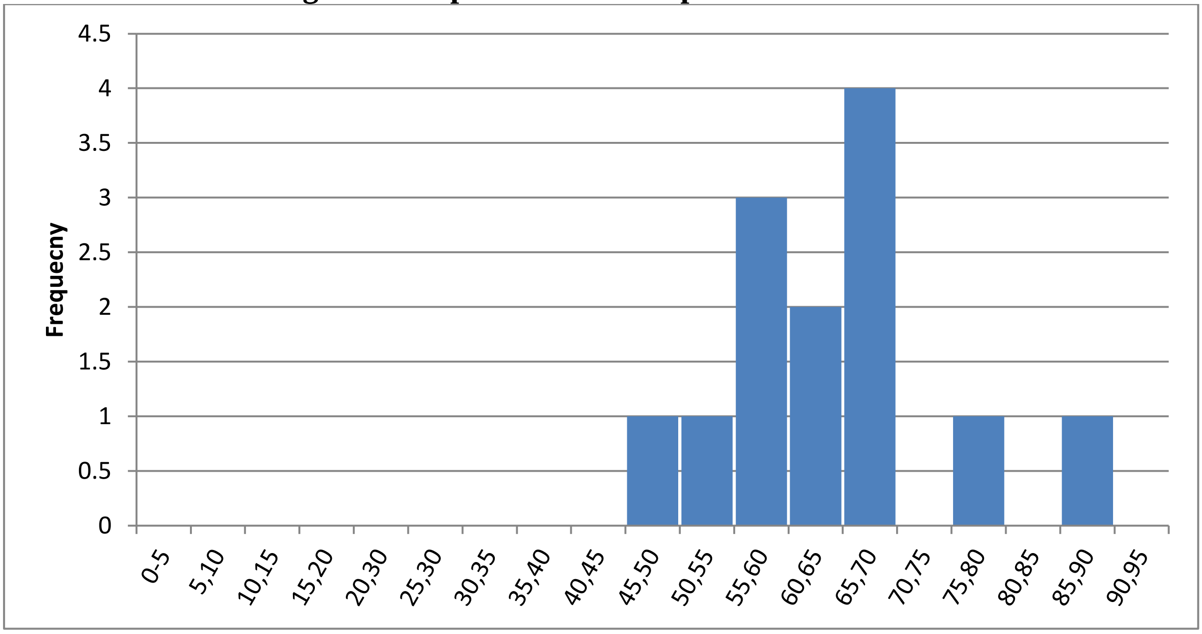


Figure 22 Experimental Group VIP Distribution



A Shapiro Wilks test shows these two distributions are normal, with p values of .29 and .81 for the control and experimental groups respectively. As seen in Table 23 below, both treatment groups results show a platykurtic distribution but the control group shows a slight negative skew and the experimental group a positive skew.

Table 21 Measure of Skewness; Visual Inspection Performance Distribution

Treatment Group	Skewness Statistic	Standard Error
Control	-0.16	0.62
Experimental	0.70	0.62

5.4.4 Visual Inspection Performance by Hazards

A bar chart for each hazard by treatment group is presented in Figure 23 below. To recap, control group results are in red and experimental in blue. No constructs are presented as the experimental design did not have the same precision of variable measurement as the kitchen study. Whilst this is a limitation, it should be borne in mind that the reason for the aviation study inclusion was about evaluating the effect of the lead investigator involvement in pedagogy, and potentially extending the applicability of SVS.

Figure 23 VIP for all Hazards & Plants

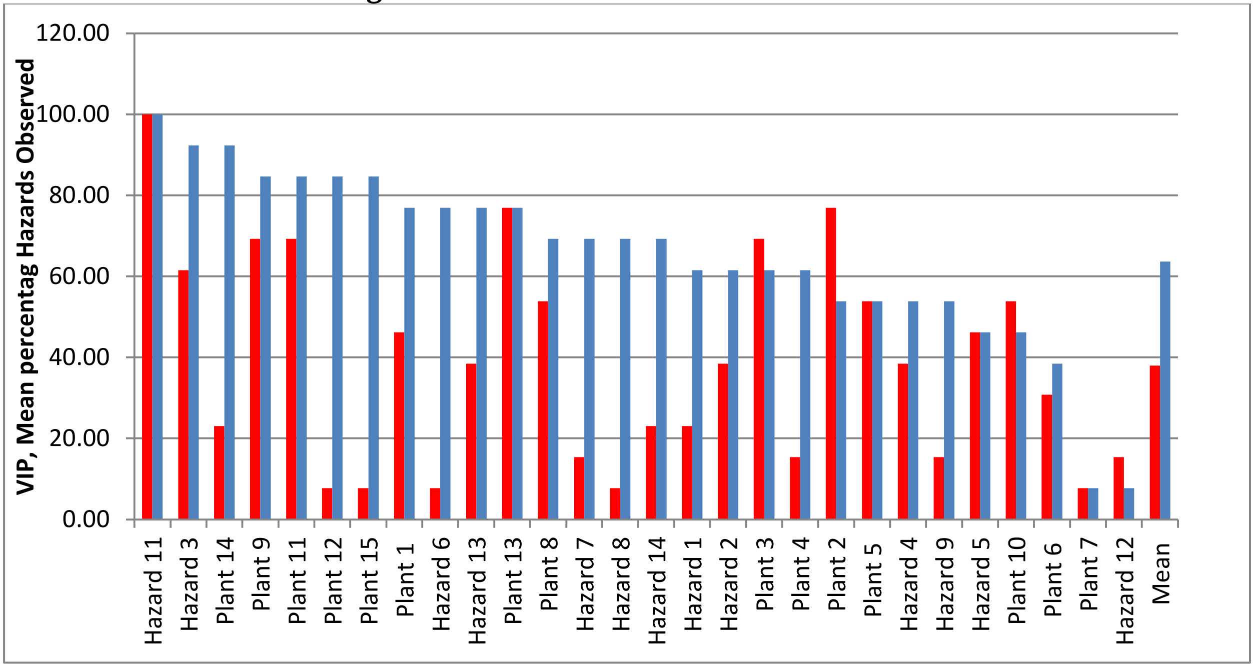


Figure 23, graphically demonstrates that SVS did not have as great an effect as was found in the kitchen study. In the aviation study, control group participants were even in VIP rates with their experimental group colleagues, in 5 out of the 29 hazard & plant categories. In 3 out of the 29 hazard and plant categories, control group participants had better VIP. Furthermore as Table 22 and Figure 23 above demonstrates, there were only 6 out of the 29 hazard and plant categories that showed a statistically significant difference in favour of experimental group participants. These were Plants 4, 12, & 14 and Hazards 6, 7, 8. Although beneficial in overall terms, SVS use by aviation participants did not have the same effect as was found in the kitchen study.

5.4.5 What Predicted Successful Observation?

With a great deal of reflection on these data, there does not appear to be any obvious characteristics of these hazard and plant categories to explain the results. Characteristics of the hazards looked at included, geographical location, whether they were located internally or externally and importance. One “plant” a flat tyre, was seen by all participants. This is very likely to be explained by the visual characteristic of size given it was the largest in terms of dimension relative to the remaining hazards which were all smaller. Size is one object attribute that is known to have a large effect on visual cognition, in terms of attentional deployment (see Table 1). With the construct of importance, 46.15% of “important” hazards were observed by control participants (SD= 18.36) and 68.53% by experimental participants (SD= 15.13). However this difference was not significant, given the conservative effect from a Bonferonni corrected p value of $\leq .0016$.

Considering the aviation study as a whole, the only factors that come to mind to explain the lack of any clear hazard attributes, are hazard characteristics and sample error. Due to experimental design and the required “hands off” approach of the lead investigator, the aviation study did not benefit from a consideration to obfuscate hazards, thereby preventing the evaluation of a variable found to affect the kitchen study. Furthermore, the sample size of 26 participants may simply have been too small to allow any “signal amongst the noise” to be demonstrated. However it remains that the use of SVS, in overall terms, did improve VIP by significant 25.73% and with a large effect size (Cohen’s d = 2.67).

5.4.6 The Number of Hazards Written Down

Table 22 demonstrates that the number of hazards written down also improved significantly with SVS users. Experimental group participants wrote down a mean 25.66 % more hazards than their control colleagues, which represented a large effect size (Cohen's *d* = 2.62). What can be seen from the histograms in Figures 24 & 25 below is that there was a negatively skewed platykurtic distribution for the control group which became a positively skewed platykurtic distribution for the experimental group. Using the same theoretical calculation as per the kitchen study of one hazard taking about 15 seconds to write down, then control group participants spent; about six and a half minutes writing, 23 and a half minutes observing and wrote nearly one hazard a minute. In contrast, experimental participants spent about 13 minutes writing, 27 minutes observing and wrote nearly two hazards per minute.

Figure 24 Distribution of N Hazards Written by Control Group Participants

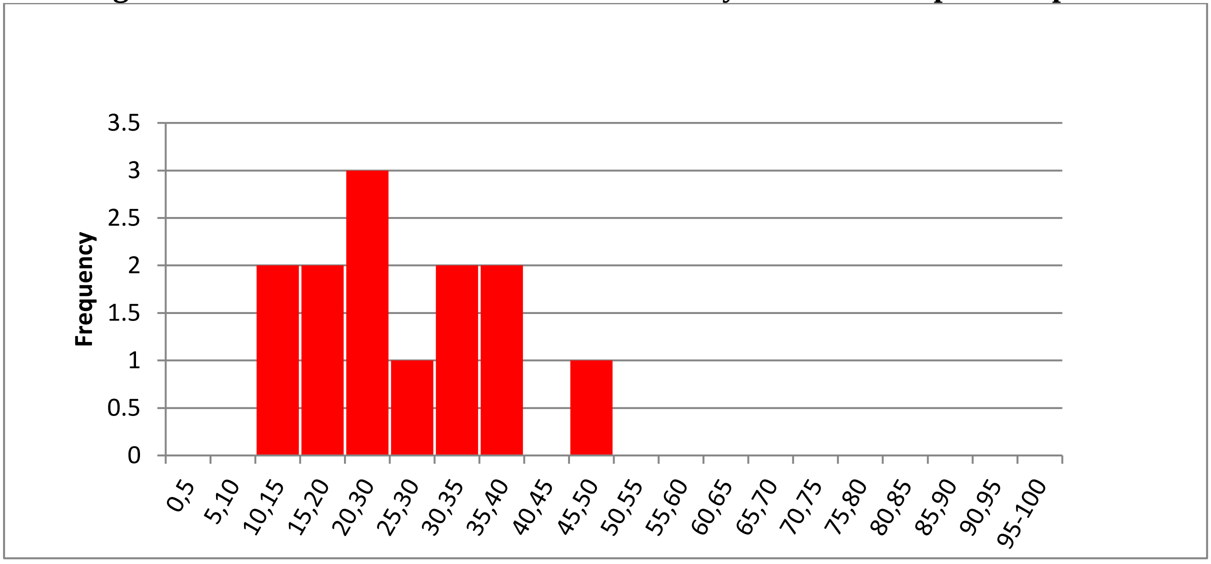
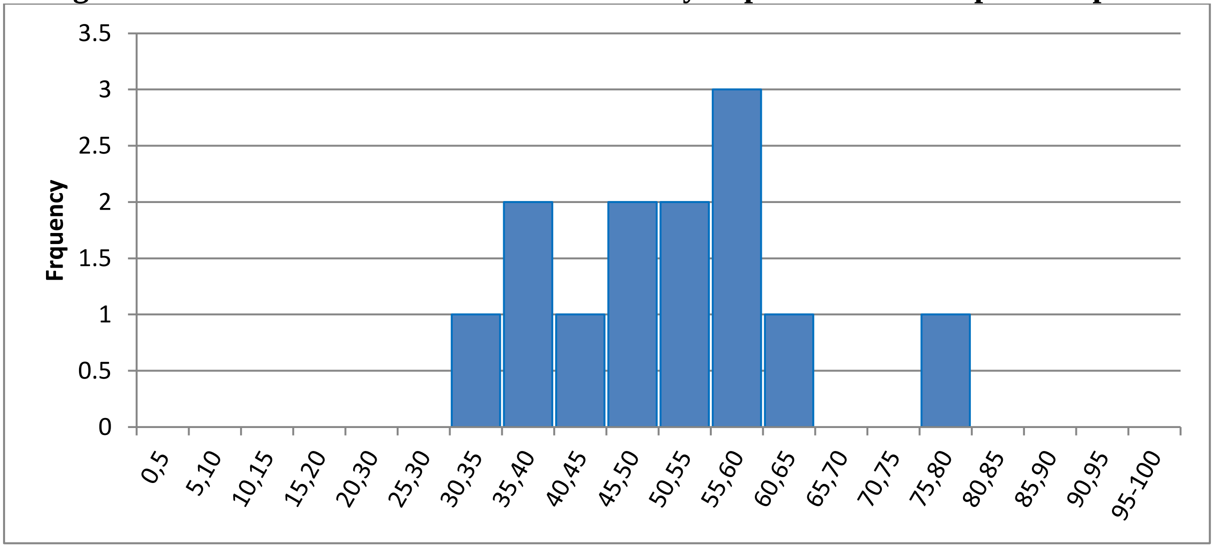


Figure 25 Distribution of N Hazards Written by Experimental Group Participants



A Shapiro Wilks test returned .63 and .55 for the control and experimental groups respectively showing normality. When described by the skewness statistic calculated by SPSS the results in Table 24 were as follows;

Table 22 Measure of Skewness; Number of Hazards Written Distribution

Treatment Group	**Skewness Statistic**	**Standard Error**
Control	0.59	0.62
Experimental	0.56	0.62

5.5 The Aviation Study; Further Inferential Results

The following will present regression and correlational results from the aviation study.

5.5.1 Regression Analysis on Visual Inspection Performance

A regression analysis was conducted to see the strength of the intervention on VIP in Table 25 below and the number of hazards written down in Table 26 below. Both these tables demonstrate the strength of the intervention's effect on VIP and the number of hazards written down.

Table 23 Regression Results for the Intervention

	b_0	SE b_0	b_1	*p*	R^2
Constant	37.93	2.70		≤.001	
BCa 95% Confidence Interval	[32.95-43.50]				
Participant Treatment group	25.73	3.90	0.81	≤.001	0.65
BCa 95% Confidence Interval	[18.57-32.92]				

5.5.2 Regression Result on the Number of Hazards Written

Table 24 Regression Results for the Intervention on N Hazards Written

	b_0	SE b_0	b_1	*p*	R^2
Constant	26.46	3.12		≤.001	
BCa 95% Confidence Interval	[21.06-32.06]				
Participant Treatment group	25.70	4.40	0.77	≤.001	0.56
BCa 95% Confidence Interval	[17.15-33.60]				

5.5.3 Effects from the Extraneous Variables Found

For the aviation study, the only extraneous variable recorded was the number of years worked by the participant. As Table 19 above on participant demographics shows, the mean years of experience by participants was very close at 24.08 (SD=6.36) for the control group and 25.08 (SD=4.86) for the experimental group. As expected, there was no effect found from this extraneous variable using ANOVA.

5.5.4 Aviation Maintenance Correlations

Table 25 Aviation Maintenance Correlations

	N Hazards Written	**N Years Working**
Control Group	r= 0.26 CI [-0.25-0.79] *P*=.39	r= -0.21 CI [-.073-0.32] *P*=.5
Experimental Group	r= 0.50 CI [-0.75-0.95] *P*=.08	r= 0.44 CI [0.085-0.74] *P*=.13

Some caution needs to be used here as none of the *p* values above show significance. Nevertheless, the only large effect seen was the relationship between VIP and the number of hazards written by experimental group participants. This efficiency effect, as with the kitchen study, can again be thought of as a speed and accuracy improvement from the use of SVS. The relationship between VIP and experience was mixed. Control group participants had a negatively correlated and low effect, whilst for experimental group participant it was positive with a low effect. There is no obvious explanation for this finding given the closeness of all the participants (see Table 21 above). The most plausible explanation here seems to be sample error, due to the small number of participants involved.

5.5.5 Summary of Main Findings; the Aviation Study

As Table 22 demonstrated, it was found that SVS significantly increased VIP from a mean of 37.93 to 63.66%, an increase of 25.73%. Table 22 also demonstrated that the effect size of this increase was large (Cohen’s *d* = 2.66). Table 22 further demonstrated that the number of hazards written also increased from a mean 26.46 to 52.15. The effect size was again large; (Cohen’s *d* = 2.62).

5.6 Comparing the Kitchen & Aviation Studies

The two thesis studies differed in a number of ways with the main contrasts being; the number or participants, statistical power, as well as the nature and type hazards being observed. A further contrast was that the kitchen study was a room whereas the aircraft study involved an object. However both studies have a number of commonalities; the same intervention, the same randomisation into two treatment groups, the same comparability between treatment group participants, the same inspection time, the same mathematical method of generating results, the same background data in terms of N hazards written and N years working and finally the same participant category; professionals, the vast majority of whom, had extensive experience. Where similar variables existed, cross study data was generated for comparison, and shown in Tables 28 & 29 below.

Table 26 Comparative Analysis of Common Descriptive Data from both Studies

Study	Control % VIP M & (SD)	Experimental % VIP M & (SD)	Cohen's *d*	Control N Hazards written down M & (SD)	Experimental N Hazards written down M & (SD)	Cohens *d*	β r^2
Kitchen VIP	32.97 (9.03)	49.64 (10.89)	1.85	31.22 (10.36)	48.86 (12.92)	1.7	0.60 0.37
Aviation VIP	37.93 (9.65)	63.66 (10.11)	2.67	26.46 (9.82)	52.15 (12.43)	2.62	0.56 0.77

Table 27 Comparative Analysis of Common Correlational Data (*r*) from both Studies

Study	Control N Hazards Written Down	Experimental N Hazards Written Down	Control N Years Working	Experimental N Years Working
Kitchen VIP	0.29** CI [0.17-0.45] *P*=.002	0.52** CI [0.39-0.64] *P*=0.00	-0.18 CI[-0.31-0.06] *P*=.06	-0.21* CI[-0.35-0.04] *P*=.03
Aviation VIP	0.26 CI [-0.25-0.79] *P*=.39	0.50 CI [-0.75-0.95] P=.08	-0.21 CI [-.073-0.32] *P*=.5	0.44 CI [0.085-0.74] *P*=.13

** p=≤.01 *p=≤.05

What is of note from Tables 28 & 29 above, is that there are five categories of variables where similar results between the kitchen and aviation study were seen namely; improvements in VIP, number of hazards written, large effect sizes, positive correlations between VIP and the number of hazards written and negative correlations with N years working. It is also of note that such similarities were demonstrated, even though the aviation study had a technical engineering focus and thereby can be described as a very different visual inspection scenario.

The first point to make is that the effects of SVS on VIP were in large part, replicated by the aviation study. This also supports the idea that bias from the role played by the lead investigators involvement in the kitchen study does not seem to have affected the outcome. This is because with the aviation study, delivery of the training session was conducted by the senior aviation engineers themselves, and not the lead investigator. Whilst it is not possible to state the lead investigator did not have a positive motivational effect on the kitchen study experimental participants, his effect will have been much lessened with the aviation study.

This similarity in so many findings by the aviation study, also points to the applicability of the SVS method not only to other industrial inspection requirements, but also from a room to an object. The kitchen study used a room with definable building elements such as ceilings and walls to apply the “reading a book” eye scanning pattern (see Figure 1). The light aircraft needed this eye scan pattern to be applied to the varying contours of the aircraft’s frame, whereby strict delineation is not as easy as a room. Yet the improvement in VIP found in the kitchen study, was largely replicated by the aviation study showing evidence for SVS’s application to further visual inspection scenarios, including objects.

However, there were some notable differences which cannot be readily explained. The main difference found was that in the aviation study, control group participants matched their experimental colleagues in 5 out of 29 hazard and plant categories and were better in 3 of them. Why this was found is open to speculation and with no obvious explanation. The aviation engineers were more experienced in term of years working, but the effect size with the use of SVS was larger so experience does not seem to be a factor. The aviation study participants also demonstrated better VIP than their comparable kitchen study experimental participants, but this can be readily explained in terms of the narrower range of hazards that would normally accompany a light aircraft defects list.

However the remaining similarities represent a good deal of replication for a field based experimental psychology based research study. Experimental bias can easily result from field based studies (Field, 2013). So to have five similar categories of results across two very different field based studies, only strengthens the experimental design and further underscores the strong effect on VIP from use of the SVS method.

In summary, VIP in both studies are broadly similar. The differences in VIP from the two control groups for each study varied only by about 5%, but rise to a 14% difference for the experimental groups. For the two control groups, this five percent difference is not a great deal and could easily be explained by internal bias. A 14% difference in the VIP between the two experimental group participants is somewhat larger, but this result by no means detracts from the similarities found in the two control group results.

This comparative analysis, especially as demonstrated by control group participants, may also suggest that visual search behaviour could well have an underlying distribution. Furthermore this distribution found in the two thesis studies could potentially be approximating   normal human visual inspection behaviour, with VIP somewhere in the thirties range percentage wise. It could be simple co-incidence that both studies in this thesis demonstrated VIP in the thirties range percentage wise. However, this idea does not seem intuitively likely. This is because there were other control group similarities as seen in Tables 28 & 28, and this VIP range has been demonstrated in four out of six studies published on VIP in the literature (see Chapter 3 (Visual Inspection Literature). If these results are generalised, some thirty to forty percent of all observable hazards could roughly be the best VIP scores achievable by EHS professionals, without any intervention during their visual inspections.

5.6.1 Overall Summary of both Studies

There were similarities in both studies that suggest SVS is applicable to different visual inspection tasks, and can be extended to objects as well as rooms. The utility of SVS seems to lie with its behavioural algorithm dimension, allowing it to be applied to such diverse observational tasks as kitchen workplaces and light aircraft. There is now some evidence that current custom and practice for VIP can follow a platykurtic distribution for SVS users. However, VIP levels were lower than expected or desirable in both treatment groups and additional research to improve visual inspection is therefore needed (see Chapter 7 Conclusions and Recommendations).

5.6.2 A Qualitative Commentary on the Results

There were some examples of qualitative commentary that were collected by the lead investigator during the kitchen study. When the lead investigator had collated scripts from experimental group participants in the kitchen study, there were a number of comments overheard and written down. One quote from a participant in the first pilot study, gave a memorable descriptive comment on how to apply SVS by stating users need to;

“paint the room with your eyes”

Other comments related to the advantages that SVS conferred;

“It shows you how to do an inspection”

“Its about time we were told how to inspect”

“That’s how I’m doing inspections from now on”

“The half hour flew by, I ran out of time and didn’t finish the inspection”

These comments allude to the utility of SVS when conducting visual inspections and seem to evidence the idiosyncratic nature of visual inspections. These comments also point towards a realisation of the advantages that SVS brings when compared to existing methods. However, these comments are anecdotal in nature and were not collected under experimental conditions. Therefore they do not have the same level of reliability and validity when compared to the interval ratio and categorical data generated in both thesis studies. Nevertheless, these subjective comments do provide an additional commentary on the effect both studies had on participants. These additional descriptions, also strongly support the need for qualitative data in any future studies. This will be further detailed in Chapter 7 (Conclusions and Recommendations).

CHAPTER 6 DISCUSSION

6.01 Introduction

The aim of this thesis was the trial of systematic visual search (SVS), a method to potentially improve the observation of workplace hazards during visual inspections. As shown in Chapter 5 (Results), this aim was successfully met. Therefore SVS has demonstrated its beneficial utility for visual inspection practice and this finding will be explored in more depth. This chapter will also consider the research question; why are visual inspections often flawed and how can these flaws be overcome?. This is far more complex when compared to the aim and hypothesis and explains the amount of content exploring this question.

This chapter opens by detailing the strengths and limitations of the two studies in this thesis. It will then compare thesis results with the visual inspection literature before speculating on how SVS may work. This will be followed by attempting to explain the generally low levels of VIP demonstrated in the two studies. The chapter concludes with a discussion on whether SVS could or should, be recommended to the wider EHS community.

6.1 Strengths of the Thesis Studies

As detailed in Chapter 5 (Results) the findings seem clear. The robust experimental design generated reliable and valid data to support the finding that SVS is beneficial when observing workplace hazards, when compared to customary visual inspection practice. The clarity of the findings can be attributed in large part, to the experimental design. The support for this argument starts with the epistemological philosophy used. Of the ontological choices possible, quantitative approach was chosen and framed the research question into a null and alternative hypothesis for subsequent evidence based testing.

The strengths of the experimental design are clear and include the following. A randomised controlled trial methodology, which is an epistemology specifically aimed at evaluating interventions (Aronson, 2012). A large sample size of 211 participants sufficient in number to achieve statistical power for the kitchen study. Randomisation of participants into comparable treatment groups was also achieved, ensuring that the intervention’s effect was able to be elicited and thereby demonstrated, by experimental group participants.

A further strength relates to the clarity of data generated. The intervention did not return closely run results with VIP differences in low single figures and *p* values hovering dubiously close to .05. Instead the two thesis studies returned loud "signal to noise" data about the intervention's beneficial effects. The data painted a very detailed picture of the beneficial effects of SVS on VIP. In addition, the use of ecologically valid conditions, allowed a realistic comparison between the simulated observational tasks in this thesis, and "real world" visual inspection practice. In short the experimental design can be considered robust and can be recommended as a methodology for further research relevant to visual inspection inquiry (See Chapter 7 Conclusions and Recommendations). The utility of combining all the dependent variables into an overall measure of VIP was also found to be very a useful construct as a dependent variable. A final strength is that certain findings from the kitchen study were replicated by the aviation study, underscoring the utility and applicability of SVS. Nevertheless, there are limitations with the two thesis studies, which will now be explained.

6.2 Limitations of the Thesis Studies

A limitation on both the kitchen and aviation studies is possible bias from the experimental design, which could have affected the reliability and validity of the data. The individual sources of bias are detailed below and will be categorised under; creation of the simulated observational tasks, task assessment & invigilation, and the lead investigator's role.

6.2.1 Bias from the Creation of the Simulated Inspection Tasks

To begin with, in both studies the random allocation sequence method used consisted of the researcher or senior aviation engineer, verbally allocating participants to treatment groups using the codes 1 or 2. With hindsight, it is accepted that this method of random allocation is outdated and not considered scientific. It may therefore have introduced some bias by not perfectly randomising participant allocation. Better methods would have included using SPSS to generate random numbers. If applied to participants in the order generated with even numbers being assigned for control groups, and odd numbers for experimental groups, then random allocation would have been achieved. (Schulz & Grimes, 2002).

It is clearly stated by some methodologists, that anything less than a strict adherence to the correct randomisation process, will bias the results (Schulz & Grimes, 2002). However, this methodological high bar has been challenged for example by Howick & Mebius, (2014), who state that there is little evidence that small deviations from strict random allocation methodology has resulted in the wrong findings from trials being published. It remains however, that participants in the kitchen and aviation studies should have been allocated to treatment groups using a more scientific random allocation method. In defence of this random allocation method, Tables 14 & 21 did not show significant differences in any demographic characteristics between treatment groups in both studies.

Bias could also have been caused by the (so called) Hawthorne effect which refers to a change in behaviour as a result of being observed (Hutchison & Styles, 2010). If this was the case, participants may not have acted as they normally would, due the presence of the lead investigator and the gatekeeper in the kitchen study, as well as the senior engineers in the aviation study. Aronson, (2012) mentions “cover stories” which are designed to increase experimental realism given that intelligent adults will try and see behind studies, which could affect their subsequent behaviour during interventions. The advice from this scholar is to use cover stories that mask the true intention of the intervention. This advice was followed by the lead investigator in the kitchen study. Randomisation was described as being necessary in order to produce two groups, thereby preventing the kitchen being overcrowded during inspections.

In terms of sample recruitment for the kitchen study, there is a geographical selection bias. The population recruited was drawn from participants attending, and thereby residing, in locations accessible to the Dublin Institute of Technology. This would preclude a geographical representation of the Irish EHS community. It can also be argued that this sample constitutes three distinct sub-populations; undergraduates, part time students and full time professionals. Regarding the aviation study, it is accepted that the sample size is too small to be representative of the wider aviation maintenance community. If this selection bias argument is accepted, then generalisation to the wider EHS community will be compromised as it represents a different population, and is therefore not comparable. Furthermore, the participants recruited for this thesis represent a “convenience” sample, rather than a random sample selected from the EHS community (Field, 2013; Tabachnick & Fidell, 2014). Therefore, there are limitations in trying to equate thesis study participants as being representative of the wider EHS community. This point is further discussed in section 6.7.1.

Whether the participants had 20/20 vision, could also not be assessed within the study resources available. However, scholars such as See, (2013) and Drury & Watson (2002) do not report inspector visual acuity performance, as being an important factor affecting results. It was therefore decided at the experimental design stage that leaving this bias in situ, was the most pragmatic solution given the resources available to the lead investigator. Furthermore the randomisation process would have equally distributed any variations in participant visual acuity.

As whole classes and years were used in this study, participants would have been inspecting with their colleagues which in many case would have been their friends. Camaraderie may have played a part if some class peers had their behaviours influence other participants. For example they may have influenced some participants to finish early, or not engage fully with the intervention. The exact opposite could also have occurred.

Another possible impact on the study is from the thirty minute visual inspection period used in all trials as there was no correction made for kitchen size, or number of existing hazards in kitchens. Other than the aviation inspection task, the number of pre-existing hazards varied and no attempt was taken to alter the inspection duration. The mean number of hazards per kitchen was 36.17 with a standard deviation of 4.04. However, the minor variations in the number of hazards present did not lead to significant variations in VIP within each treatment group. Therefore, this bias effect is considered to be limited in nature.

Regarding the kitchen study, there was also the possibility of participant contamination. To explain; participants in earlier trials may have come into contact with participants from later trials and informed them of what to expect. This could have led to later trial participants modifying their behaviour due to prior knowledge. Whilst this remains a possibility, the probability of this having occurred is considered limited to due to how participants were recruited.

In terms of the two trials that involved full time workers, it is unlikely they would have had contact with each other as they were from different organisations. Similarly the part time and full time students were on different educational programmes again with no direct interaction. It is possible that participants may have known each other socially and thereby discussed the trial procedures. However as Figure 20 shows, the results of VIP performance by trial do not show any incremental improvement which would evidence if later trial participants had advance knowledge of the experiment.

Finally, multiple participants all visually inspecting at the same time is not realistic of industry practice and therefore impacts on ecological validity. However it was logistically not possible to allow 237 participants to conduct 30 minute inspections sequentially and on their own. But it remains possible that VIP could have varied if inspections were conducted individually.

6.2.2 Bias from Task Assessment & Invigilation

The lead investigator was aware from the pilot studies that a certain degree of ambiguity was to be expected from the scripts. This was minimised by applying a consistent diligence in the marking. Blind marking was considered at the experimental design stage. However, this was not possible due different paperwork being supplied to each treatment group for recording observations. Experimental group participants received paperwork with a pre-ordered selection of elements to facilitate SVS (see Appendix 4). In contrast control participants wrote their observed hazards on A4 paper they brought with them into class. Therefore it was not possible to use the same paperwork in the interest if blind marking for all participants. This difference in paperwork used, can be argued to be weakening comparability between participant groups. However it is intuitively difficult to see how the paper used to write down observed hazards, would have biased VIP in control or experimental groups to the point of affecting the results.

The lead investigator may have been on a very small number of occasions, overly corrective when marking participant scripts. It is conceivable that if the marking was more forgiving or allowed far more credit to any of the very vague written notes encountered in the scripts, then VIP would have increased. Furthermore, it would be difficult in the absence of an inter-rater analysis, to exactly measure the hazard observation ability of 263 participants from their scripts. Even if inter-raters were used, there would still have been issues in the interpretation of the scripts as there were numerous examples of ambiguous hazard identification. For example, participants writing down; “food left of the floor” illustrates this ambiguity. Does this phrase count as the observation of one food item on the floor or more. If more than one, was it two, three or four. This was dealt with in the following manner. In order to recognise hazard observation, even when ambiguously recorded, the lead investigator chose to acknowledge rather than ignore such ambiguity in written accounts of hazard observation. So, a vague phrasing such as “food left on the floor” was interpreted in the plural as two hazards visually identified. In contrast, if a box of food on the floor was written, this signified one hazard observed. For more than two food items on the floor to be identified, the participant would have had to write three such observations separately on their script. The same situation and interpretation was applied to phrases such as “bottles of chemicals left in the kitchen”.

A particular problem in the kitchen study, was categorising vaguely phrased general hygiene issues. Participants were writing various descriptions of unhygienic conditions for example, “dirty plates were left in the sink, dish cloths were dirty, utensils had food remains, the kitchen needed a thorough cleaning” etc. This was a particularly intractable problem as the number of hazards written down by participants, in many cases exceeded the actual number of known hazards within the trials. It was felt necessary to capture this information as it did represent participants identifying unhygienic conditions and furthermore, some participants did not record any observations of cleanliness. This was again dealt with by choosing to acknowledge this type of observation rather than ignoring it. As a result, there was one variable created to cover participants observing, or not observing, overall hygiene conditions. Individual variables were further created to record particular hygiene conditions for example floor cleanliness, north wall cleanliness, or south wall cleanliness, which then represented three distinct hazards. Therefore a common participant response of; the kitchen needs a thorough clean, and the floor near the sink is dirty, would be regarded as two hygiene hazards identified.

The orientation or physical location of the hazard within the kitchen also proved problematical again due to participant phrasing. For example if a participant wrote; a wall was dirty, it was not possible to locate the particular wall referred to. Once again, this was dealt with by recording such phrases as a hazard observed. However its location was entered as not applicable by the lead investigator.

Due to the nature of the hazards themselves, it was not possible to see if there was a particular visual attribute that could predict observation. This was due to the difficulty in describing salience or terms such as conspicuity, “noticeabilty” or how much a hazard stood out from its surroundings. In practice, to create a variable that could describe the differences in salience or noticeability would have been far too subjective to have empirical reliability and validity. A measure of perceptual sensitivity to salience that has been used is the “d” index, measured by reference to a signal to noise approach (See, 2012). This index has been used to describe inspection tasks from very easy, moderately easy, moderately difficult and very difficult. However this index as cited by See, (2012) was reported by Craig, (1984) but has not been prominent in the visual search literature since that study, and was therefore not used in this thesis.

It was also possible that participants copied their results from other participants. Table 12 shows that the total number of participants per trial varied from five to eighteen. When the treatment group sizes were greater, the opportunity for the plagiarism must have been present. It was also possible that a form of plagiarism occurred, when one participant manipulated the environment allowing another participant to view. For example, one participant may have opened a fridge door allowing another participant to view hazards, they would not have otherwise observed.

It was also possible that error occurred in transferring data from scripts into electronic formats. Transfer was checked after entry by the lead investigator, and descriptive statistical tests run to see if any anomalies were present. Although none were found, the possibility of error remained given that over 22,000 data points were entered electronically by the lead investigator. Appendix 5 lists those hazards that constituted individual variables that were amalgamated into separate constructs such as visual inspection performance (VIP). Inherent data entry errors would have been transferred into the new categories, and therefore would still be affecting accuracy. However this would be mitigated, as a greater number of data points were used in these larger categories, which in turn allowed for greater descriptive and inferential statistical confidence in the results.

6.2.3 Bias from the Lead Investigator’s Role

In addition to the intervention there were two further, but necessary differences that weakened slightly the exact comparability between control and experimental participants. The first was that SVS training meant an extra half an hour of contact time with the lead investigator, by experimental group participants. The second difference was that the invigilator for the visual inspections for control participants (the gatekeeper) was a different lecturer. This was necessary because the lead investigator was delivering the SVS training at the same time as control participants were visually inspecting.

So by having an extra half hour with the experimental group participants, the lead investigator could have had a motivational effect resulting in these participants trying harder than their control group colleagues. Although this bias is possible, the aviation study demonstrated broadly similar results without his involvement. What can be logically said here is that the extra half hour with the lead investigator cannot be discounted from affecting the kitchen study participants. But, considering that the kitchen study results were broadly replicated by the aviation study, this influence does not seem to have manifested itself to any great extent. It is also possible that the control group invigilator could have had a motivational effect on her participants. Again this is difficult to quantify, but given the invigilator was also a lecturer of similar standing and experience to the lead investigator, this bias would also seem to be limited in effect.

The lead investigator is also just as susceptible to conscious and unconscious bias as any other individual. Therefore the results, especially in the kitchen study which was conducted by the lead investigator, may have been influenced by this bias to favour SVS. This has been discussed by the prestigious Journal; Science (Open Science Collaboration, 2015). Here, the authors reproduced 100 psychology based experimental and correlational studies published in 2008. They found that when replicated, only half of these studies presented similar results to the original studies. Whether bias or deliberate data manipulation to favour SVS took place in the study is difficult to prove or deny. Other than confidence in the fidelity of the lead investigator this study, as in any other experimental psychology study, greatly benefits from replication, to support the findings presented in this thesis.

In summary, bias from the experimental design, as in all field based studies was expected to be present to some extent, but it is difficult to evaluate its exact extent. However, any internal bias should be viewed against the robust experimental design. The clarity of data also allowed significance testing, effect sizes and inferential statistics to be assiduously applied throughout the thesis to support confidence in the results. So after much reflection it is argued that the largest source of bias that may have affected the data generated, was probably from ambiguity in marking the scripts and which may have resulted in a small number of false positives and negatives in the recording of observed hazards. Even so, this small effect would have been evenly distributed between treatment groups.

The experimental group participants specifically, may have been subject to bias due to the lead investigator’s extra half hour spent with these individuals. This may have introduced an extra motivational effect for experimental group participants only. However, there is little evidence to show that this extra half hour contact had any tangible effects due to the aviation study results. Therefore in overall terms, the robust experimental design, the comparability of treatment group participants, the clarity of the results and the large effect sizes do not suggest that internal bias can weaken the main findings to any great extent.

6.2.4 Limitations From the Thesis Experimental Design

Even though the experimental design was robust and clearly demonstrated the beneficial effects of SVS, possible improvements for future inquiry have been identified (see Chapter 7 Conclusions & Recommendations) and will now be detailed. This should not be seen as a critique of the experimental design. Instead, this is an opportunity to improve future visual inspection inquiry by using experiential hindsight gained from conducting this PhD level research. The improvements begin with the scope of thesis aim which to recap is; that SVS can potentially improve the observation of workplace hazards during visual inspections. This aim can now be viewed as somewhat narrow in outlook as more visual search inquiry could have been conducted. For example, the lack of within group trial repeats, whereby control and experimental participants could practice their visual inspections and receive feedback is an experimental design limitation. Had the trials been repeated using feedback, even with a smaller sample, then the effect of practice and informing participants of what they correctly observed and what they missed, would have been most useful. In particular, control group participants, having been told of the hazards they missed simply because they didn’t look in any obfuscated areas, could have altered their subsequent visual search behaviour to go looking at those areas they previously didn’t. In this way, they may have matched or even exceeded the VIP of their experimental group colleagues.

A further experimental design limitation was constantly having to ensure that participants did not see through the experiment, especially when they were randomly allocated into treatment groups. The pilots showed that random allocation had to be done quickly, to prevent any possible second guessing by participants that they were going to be treated differently. A corollary to this is; once the trials were conducted, further trials would be viewed suspiciously and participants would have time to talk to each other. This means that any pre-post experimental design or subsequent intervention has to take this suspicion into account.

Another experimental design limitation is the difficulty in attributing effects to the intervention when this intervention itself, consisted of three separate components. These components were; the efficacy of the SVS intervention, the quality of training and materials (including different invigilators and paperwork) and the effectiveness of execution of SVS by individual participants. Together these three components represented the intervention, but the experimental design did not address the proportional effects of each component. Again this will need further research to determine if any of these three components are independent variables in themselves.

The use of qualitative inquiry would also have been most beneficial. For control group participants, questions as to how they conducted their visual inspections would have yielded valuable insights into their visual search strategies. Evidence for hybrid foraging visual search behaviour or whether cognitive bias was occurring, could also have been elicited using semi structured interviews. For experimental group participants, questions on how they conducted their visual searches could have further supported SVS by listing the advantages and disadvantages of this method. The main reasons for not implementing a qualitative strand to this study, was simply pragmatic. The lead investigator as a lone researcher could only conduct so much fieldwork with the resources at his disposal. For a start, it was not possible to know in advance how many interviews would be needed to ensure reliability and validity as is expected in PhD qualitative inquiry practice. Interviews would in themselves have been very difficult to organise given that only a two hour period was available to participants for the trials, so questionnaires would have had to be employed.

Furthermore, for something as complex as visual search behaviour, interviews would have been preferable to questionnaires. Also the use of qualitative commentary would have moved away from the thesis aim which to recap was; the trial of SVS, a method to potentially improve the observation of workplace hazards during visual inspections. When the experimental design was being considered, empirical VIP data was a primary objective and (unfortunately) qualitative commentary did not have the same impetus. In the end, the gathering of enough qualitative data with sufficient reliability and validity was considered beyond the resources available to the lead investigator.

As a counter to these experimental design limitations, it can be argued that it is very rare that one study, especially a field based study on the scale used in this thesis, can satisfy all possible angles that an aim, hypothesis and question require. In addition, a further positive aspect to the experimental design used is that whilst some questions were answered, these very answers prompted more questions (see Chapter 7 Conclusions and Recommendations). This of itself, suggests a measure of success from this thesis that can support the need for further visual inspection inquiry in the interest of workplace safety.

Further limitations of the systematic visual search method relates to non-applicability to complex processes in very specific work environments or the evaluation of risk in such scenarios. The phrase complex work processes is here presented here in a literal sense and described as systems that contain a high degree of complexity whereby visual inspection is limited in diagnostic terms. For example Demichela & Camuncoli (2014) and Gerbec et al, 2017). give examples of complex process as; an industrial processing plant that uses the chlorination of propylene at high temperatures to produce allyl chloride, nuclear power plants or large scale industrial gas storage.

Similarly systematic visual search with its emphasis on the optimising the observation of workplace hazards, has a limited role when quantifying the risk from identified hazards. This limitation ensures that the overall evaluation of hazards, is not a process that systematic visual search can readily assist with. However the quantification of risk is a field of inquiry with a well developed literature, and with many mathematical methods (for example see Gerbec et al, 2017). Furthermore, with the quantification of risk, the analyst can rank the identified risks and assign descriptors such as low or high risk (Demichela & Camuncoli, 2014; Gerbec et al, 2017). Again the decision to classify a risk into low or high is not a function of systematic visual search.

6.3 Relating the Thesis Findings to the Literature

Due to similarities in results, there are visual search studies that are supported by the kitchen and aviation study findings. To begin with, a sentiment that is expressed by various scholars and demonstrated by the thesis findings, is that consistency in visual search behaviour is beneficial for VIP (Biggs & Mitroff, (2013b); Biggs & Mitroff (2014) Nalangula et al (2006); Nickles et al, (2003) Melloy et al (2000); See, (2012); Wang et al, (1997).

Visual inspection has been described by various scholars as a difficult task to do well (Biggs and Mitroff, (2013); Cain et al, (2011); Drury & Watson (2002); Gallwey, (2006a); Rao et al (2006); See, (2012) Wolfe, (2005). The generally low VIP rates demonstrated by the largely professional kitchen and aviation study participants also reflect this sentiment. The two thesis study results also echo the overarching comment from See, (2013) when she states that human visual inspectors are imperfect, and exhibit large differences in performance across a range of inspection tasks.

The findings from this thesis are also in broad agreement with the six field based studies that most closely resemble the experimental design used in this thesis (see section 3.1.3) In particular this thesis reinforces the Albert et al, (2014 & 2017) studies from US construction sites that found current custom and practice for VIP, was in the thirties region percentage wise.

Another similarity between the literature and this thesis, is how VIP becomes somewhat higher when there is a lesser number of potential hazards to be observed. Section 3.4 presents VIP data from the industrial quality, security screening and medical diagnostics sector, which all demonstrate higher VIP when there are a lesser number of hazards in the area under analysis. In this regard, the thesis aviation study is generally in keeping with this literature. In this thesis, there was little evidence that salience was a predictor of the observability of hazards by participants. This has also been reported in the visual search literature (see for example Biggs & Mitroff, (2014b).

The literature and thesis findings both have mixed results regarding VIP and experience. Some studies have shown that experience does have a positive effect on VIP (Biggs et al, 2013b; Biggs & Mitroff, 2014b; Wales et al, (2005). However, Bahn, (2013) found that experience was not a predictor of hazard identification ability. This mixed picture found in was also found in the two thesis studies (see sections 5.2.8 & 5.5.4).

6.4 Evidence for How Systematic Visual Search Works

This section will now present a discussion that could potentially explain how and why, SVS was found to be beneficial when compared to control group participant VIP. In the first instance, salience was not found to be a predictor of observation, and thereby can be eliminated from explaining any SVS mechanisms. The next part of this discussion, presents evidence from the two thesis studies and literature that suggests the beneficial effects of SVS could have occurred from two processes. Firstly, SVS is a behavioural visual search algorithm which promotes the exhaustive search of the area under analysis, leading to the better deployment of visual attention. Secondly, the literature theorises that SVS is aided and abetted by the observer using minimum cognitive resources, at maximum efficiency. These explanations are now explored in detail.

6.4.1 A Lack of Salience Characteristics Predicting Hazard Observation

The kitchen study control participants demonstrated that the only hazard characteristic that reliably predicted observation, was whether or not the hazard had a direct line of sight by being in open view. Only in one case did salience, or how much the hazard stood out from its background, seem to manifest itself. All participants in the aviation study noted the flat tyre on the light aircraft.

However, the only statistically significant hazard characteristic to emerge was whether or not there was a direct line of sight to the hazard in the kitchen study (see section 5.1.5). It should be remembered that the hazards themselves were not technical in nature, many of them were in plain sight and for want of a better description were visually obvious. For the kitchen study, they included pictogram labelled bottles, cigarettes in ash trays, cooked pies left on shelves, packets of food on the floor, fire doors propped open with fire extinguishers. For the aviation study they included mobile phones, coins, pens and screwdrivers left unsecured. These hazards can all be described as items that were not particularly difficult to observe. For example, the skull and crossbones pictogram labelled bottles of water did not show any real difference in the rate of observation from packets of food left on the floor. Assuming there is no salience effect, the question becomes; how did SVS achieve its beneficial effect in seeing more observable workplace hazards?

6.4.2 SVS Promotes Exhaustive Search & Better Attentional Deployment

There is evidence from the two thesis studies, that SVS achieves its beneficial VIP effect by better promoting exhaustive visual search, which in turns guides better visual attention deployment for observable hazards. The argument and the evidence for this, is as follows. In the kitchen study, Figure 13 and Table 15 demonstrated that SVS users observed 12.54% more hazards in open view than their control colleagues. This finding was statistically significant and with a large effect size (p=≤.0017; Cohen's d = 1.05).

Furthermore, SVS users observed more hazards located on the ceiling and in obfuscated areas of the kitchen, as demonstrated in Figure 13 and Table 15. Therefore, SVS users looked far more often, at areas that their control group colleagues simply did not look at in such detail. In short, SVS users visually searched more areas and they searched these areas more thoroughly. The evidence for searching more exhaustively is compelling.

If an improved thoroughness of visual search is a direct consequence of SVS use, then the corollary to this suggests that VIP improvement is fundamentally achieved by more instances of fixating on observable hazards, which in turn better guides visual attention deployment for those observed hazards. In effect by searching more thoroughly, SVS users were more likely to deploy their visual attention when guided by bottom up stimuli processing to better see any observable hazards present.

In addition, the use of foveal vision could also have an important role here. To recap the SVS training emphasised the use of foveal vision by requiring participants to locate themselves as closely as possible to the area under analysis. Foveal vision has the greatest object resolution capability which if maximised, could also have played its part by promoting the better visual resolution of fixated hazards and thereby better visual attention deployment (Gallwey, 1998a; Wade and Swanston, 2013).

In summary, the overall effect that SVS potentially brings can be now succinctly encapsulated by using a chain of events description. It begins with SVS promoting a more thorough search of the area under analysis. This in turn makes it more likely for visual inspectors to “clap their eyes” on observable hazards present. This in turn results in better deployment of visual attention with a conscious perception and understanding of the actual hazard seen. Finally this perception and understanding allows for a cognitive decision analysis process resulting in the hazard being written down. In contrast, control group participants using their customary practice of visual inspection, did not search the room under analysis as exhaustively. As a result, they encountered less observable hazards to deploy their visual attention. Their VIP was duly demonstrated to be significantly lower.

A further question now arises. What is the potential utility of SVS, if EHS professionals simply practice and receive feedback on their customary and idiosyncratic visual search methods. Put another way; consider the scenario whereby the control group participants in the kitchen study, practiced their normal visual inspection methods ten times over, with detailed feedback given each time. On their eleventh visual inspection, would these same participants still show an improved VIP, if they used the SVS method.

Again, this is difficult question to answer with any confidence. But there is evidence from the two thesis studies that there would still be benefit from using SVS, even after repeated feedback and practice with their customary visual inspection practice. This will now be explained. Firstly, the aviation study control participants had over twenty years of experience. They cannot be described as lacking in practice. Furthermore as described by Drury & Watson, (2002), aviation is a highly regulated industry. It does not seem plausible that the aviation study participants would not have their work checked on a regular basis by their supervisors during their long careers. In effect, it is likely that the aviation study control participants were very much practiced and supervised in their usual visual inspection conduct. Yet SVS users in this aviation study still demonstrated a significant 25.73% increase in VIP (p=≤.0016; Cohen's d = 2.67).

Secondly, if VIP for control participants who receive practice and feedback improves, then there is no reason why the VIP of SVS users should also, not improve with feedback and practice pedagogies. Therefore if SVS novices also practiced their newly acquired visual search skill ten times over, it is likely they will improve as well.

Finally and perhaps most convincingly, it is the reliance on bottom up visual stimuli processing that SVS seems to exploit. SVS does not seem to require any top down stimuli such as training, or increasing attentional set. If observable hazards are present, then an exhaustive search of the area under analysis will deploy visual attention. If this is the case as seems likely from the two thesis studies, then it makes rather obvious sense to assume an efficient way of observing workplace hazards is to look exhaustively and allow bottom up stimuli processing to guide visual attention deployment. In layman's terms, if you use a set eye scanning pattern to make sure you look everywhere, then you will eventually "clap eyes" on all the observable hazards present.

6.4.3 SVS Conserves Cognitive Resources

The evidence for claiming that SVS promotes a minimum of cognitive resources used at maximum potential comes from the literature. Various scholars argue that consistency in visual search strategy, confers cognitive advantages including more efficient use of working memory (Cain et al, 2013 and Mitroff & Biggs, 2013). This effect is from the visual searcher not observing areas already looked at. Some scholars theorise that this is due to short term memory not being that capable of accurately distinguishing previously searched areas. This suggests that SVS allows users to better remember where they have searched (Cain et al, 2013 and Mitroff & Biggs, 2013). So when experimental group participants wrote down their observed hazards, the location was better retained and search resumed from this location, and not any previously searched location. This in turn prevented any unnecessary additional visual searches, and thereby conserved cognitive resources.

A good analogy would be reading a book. If a reader is distracted for example by a phone call, making a memory note of the page allows the reader to quickly return to where they stopped reading. If this memory note is not made, then the reader has to re-read a number of lines of text before finding their place on the page which uses up cognitive resources. Applying this theoretical explanation, it seems plausible that SVS could be acting in the same manner. So it may be tentatively possible to state that SVS maximises observer cognitive resources as follows. The eye scanning pattern used to observe for example a kitchen wall, ensures that this sensory data is held in the visual working memory. In effect a mental note is made by the observer preventing this element from being un-necessarily visually searched again. Furthermore, it is known that visual inspection is a cognitively demanding task (Biggs and Mitroff, 2013; Drury & Watson 2002; Gallwey, 2006a; Rao et al 2006; Wolfe, 2005: See, 2012). So it can be speculatively proposed that those observers who do not use SVS, will be visually searching areas already observed, and thereby wasting more of their available cognitive reserves.

6.4.4 Effect of Experience on Visual Inspection Performance

When VIP was correlated with the number of years worked, treatment groups did not demonstrate any consistent relationships. In the kitchen study, all participants had a negatively correlated low to medium strength relationship with experience but only experimental participants returned a significant effect ($p=\leq.05$). In the aviation study, only control group participants returned a non significant negatively correlated low to medium strength relationship with experience. In contrast experimental participants in the aviation study returned a non significant positively correlated low to medium strength relationship with experience.

Taking the kitchen study first, why both control and experimental participants should get worse at visual inspection with more experience is a counterintuitive finding. However this could be explained by a number of factors. It could be down to a form of over-zealousness with younger professionals seeing far more items as hazardous. With experience, they may come to realise these same hazards as safe. Conversely if there are no adverse effects from their observed hazards over the years, then there could be an outcome bias effect. This could result in participants normalising the abnormal, and not regarding these hazards as actual hazards. Regarding the aviation study results, there was no consistent correlations in both treatment groups, which seems to be best explained by error from the small sample size.

6.4.5 Correlating VIP with the Number of Hazards Written

The correlation between VIP and the number of hazards written down by all participants was significant (*p*=≤.01), and thereby pointed to a clearer picture. Here it was found that control group participants from both studies, had a positively correlated low to medium strength relationship between VIP and the number of hazards written down. The strength of the relationship increased to a strong relationship with experimental group participants in both studies, (however the correlation was not significant in the aviation study). This is further evidence of an improved efficiency effect from SVS in the kitchen study which was again significant (*p*=≤.0017). Here SVS users wrote down more hazards when compared to control participants. Whilst this fits in with the general finding that SVS users were better at observing hazards, it nevertheless strengthens the finding that the intervention brings an efficiency effect manifesting itself as more hazards being written down when compared to their control colleagues.

6.4.6 Strengths of Systematic Visual Search.

Section 6.1 has discussed the strengths of the overall thesis and its experimental design. It is now time to focus in on the SVS method itself. Isolating the SVS method and results from the wider thesis, allows a deeper discussion as to the strengths of this behavioural algorithm for visual inspections. One undoubted strength of SVS and perhaps its greatest attribute is its dimension as a visual search behavioural algorithm. SVS represents a stand alone method that does not seem to be reliant on any specific top down guidance such as training, or require any environmental changes to the area under analysis to make observation easier. Also, the set eye scanning pattern it promotes has demonstrated a specific type of search consistency that is beneficial for VIP. Of the methods to improve visual inspection, it remains the only one that can be used as a stand alone method applicable to all situations. It can also be used in conjunction with all alternative training syllabi and any environmental alterations.

As seen in Chapter 2 (Theoretical Underpinnings) there are three basic ways of improving VIP; improving attentional set, improving search behaviour or manipulating the environment to make observation easier (Drury & Watson, 2002; Nickles et al. 2003; Rao et al. 2006; Schwaninger, 2005; Wales et al, 2009). The problem with improving attentional set is which syllabi and pedagogy to select as there are many choices. The problem with improving the task environment to suit the visual inspector is time, cost and effort. Furthermore, this would not be possible in many cases as EHS professionals will visit a wide variety of workplaces. Finally, the problem with eye scanning patterns is again which one to use. The three published studies on set eye scanning patterns all had different search patterns (see section 3.3 for details). Simply put, SVS is the first to offer the complete package for visual inspection; what elements to look at, what order to look at them, the eye scanning pattern to use and an evidence base for its utility.

Another strength of SVS is the ease of training for the method. The lead investigator only needed thirty minutes of training time, a PowerPoint display and a nine page handout for participants. With the aviation study, the lead investigator played no hand act or part in the training, other than briefing the senior aviation engineer on the method. After that, all contact was suspended until the results were received. This reinforces the simplicity of the pedagogy involved.

A further strength of SVS is its applicability. It has been applied in this thesis, to a room (a commercial kitchen) and an object (a light aircraft). Therefore this behavioural algorithm does not rely on any further attentional set, alternative search strategy or the need to alter the work environment under analysis. It therefore represents a fundamental way to observe hazards during inspections that is not reliant on further requirements other than a pair of eyes, a pen and paper. There seems to be few visual inspection tasks that could be outside its capability. In this sense SVS has a fundamental and ubiquitous applicability that stems from its unique dimension as a behavioural visual search algorithm. It seems to be potentially universal in applicability for observational tasks where a certain level of accuracy is required. There is also some evidence that the method exceeds expectation. It is recommended by some scholars that search environments are manipulated, so that no more than six defects, hazards, or targets need observing. This is because any more than six targets, has been demonstrated to make the visual search difficult Galwey, (1998a). As demonstrated in this thesis, SVS seems to responds to bottom up stimuli processing and is therefore not seemingly constrained by the number of defects, hazards or targets present in the area, or on the object under analysis.

SVS seems to have further efficiency effects in comparison to customary visual inspection performance. This was demonstrated by experimental group participants from the kitchen and aviation studies who wrote down circa 18% and 26% more hazards respectively than their control group colleagues. In writing down more hazards they would have had less time to visually inspect. Therefore SVS resulted in a better use of the 30 minute inspection time available, with more hazards being seen and in less time. Again this evidence for this efficiency effect is compelling given the thesis data. This efficiency effect has not been reported in the visual search literature. Decreased visual search accuracy under time pressure has been reported in lab experiments (Fleck et al, 2010). It is also reported that nothing has yet been found that eliminates search error when the searcher is under any type of time pressure (Cain et al. 2013). In contrast SVS has demonstrated how less time can be used to observe and write down more hazards, and still outperform current visual inspection custom and practice.

In summary, SVS serves as a basic skill applicable to all EHS professionals. SVS can also serve to proceduralise that most fundamental EHS professional role; preventing accidents and ill health, by identifying and thereby managing workplace hazards. SVS fulfils this role by using the most economic, adaptable and useful sensory mechanism available to EHS professionals; their own eyes. In standardising visual inspection conduct by promoting SVS, EHS professionals will have a skill applicable for all workplace visits, regardless of the aims or provenance of that visit. It is a skill that can also reduce the ambiguity surrounding the conduct of visual inspections, leaving stakeholders with little doubt as to the professional standard involved during visual inspections. Furthermore, SVS would be most useful during judicial proceedings whereby the conduct of the visual inspector is under scrutiny. Using SVS and recording visual inspection details will evidence to the courts; where the EHS professional looked, what they looked for, for how long did they look and finally what they observed or did not observe.

6.5 So Why Was Visual Inspection Below Expectation?

Societal expectations of professional searchers is such that the highest standards are taken for granted (Mitroff & Biggs (2014); Wolfe et al (2005). However, whilst SVS did improve VIP the overall results remains a matter of concern from an EHS professional practice point of view. To recap; control group participants from the kitchen and aviation studies demonstrated VIP in the region of 32% and 38% respectively. Experimental participant results were in the region of 50% and 64% respectively. By any stretch of the imagination, these results are below expectation given the largely professional background of the participants, and the question that now arises is why.

This is a difficult question to answer as there are so many applicable causes of visual search error that could have contributed to this finding (see Table 3, & sections 2.3-2.6). It is further complicated by the experimental group participants who by following the SVS method, will have used a different visual search behaviour requiring its own interrogation and analysis. So to answer this question, a reflection on each individual visual search error listed in Table 3 was undertaken as to their possible role in explaining the thesis results. The wider visual search literature and in particular the nine published studies that most closely resemble the experimental design used in this thesis, were also scrutinised for possible clues to answer this question.

To begin with, there are possible causes that can be safely eliminated from consideration. The first is experimental design bias. For the reasons described above in section 6.1 above, there was little evidence that this internal bias could account for such low VIP results. A second explanation that should also be eliminated is a lack of knowledge amongst the participants. This attentional set issue seems to be an unconvincing argument. This is because all of the participants would have possessed extensive first hand knowledge of relevant hazards. In the kitchen study, all the participants were attending a third level college and studying for safety & food related qualifications, or were food service industry professionals. Similarly aviation maintenance engineers are required by law, to be highly educated, and qualified in their field. In addition they all had in the region of 25 years of experience. Furthermore both the kitchen and aviation study were accompanied by recapping lectures whereby all the hazards to be observed in the inspection tasks, were specifically presented to participants.

Hazard salience or a lack of hazard “noticeability” can also be eliminated, as discussed in sections 5.1.5. The randomised controlled trial methodology also eliminates any possible effect from an uneven allocation of participant characteristics such as motivation or visual acuity. Furthermore, the time available for visual inspections seems appropriate. In addition, there was a sufficient number of hazards within the kitchen and aviation studies available for observation. Keeping in mind the long list of causes of visual search error in Table 3, the following explanations are considered the most convincing to explain the low levels of VIP demonstrated.

6.5.1 Normal Human Visual Search Capability

It seems plausible, if not probable, that the normal limitations of human visual search capability have been demonstrated under the ecologically valid experimental conditions in the two thesis studies. The wider literature has many scholars echoing the description of visual inspection, as an error prone task that is difficult to do well. (Biggs and Mitroff, 2013; Cain et al, 2011; Drury & Watson 2002; Gallwey, 2006a; Rao et al 2006; See, 2012 Wolfe, 2005). Therefore it can be argued that it should not come as a surprise that VIP for the kitchen study control group participants, was in the thirties region percentage wise (as also found by Albert et al, (2014 & 2017).

The evidence for this is strengthened by those scholars who also report human variability in visual inspections performance (Drury & Watson, 2002; Gallwey, 1998a; Jiang et al 2000; See, 2013; Solman, 2014). This variability is also a thesis finding, demonstrated by the percentage density function histograms presented in Figures 11, 12, 15 & 16. In short, the literature strongly suggests that our limited observational accuracy during visual inspections is a basic human characteristic.

6.5.2 Cognitive Bias

Cognitive bias could also be contributing to the low levels of VIP demonstrated in this thesis. Cognitive bias has been shown to be a pervasive and well known cause of judgement error amongst decision makers (Aronson, 2013; Gilovitch, et al, 2013; Kahneman, 2011; Montibeller & Winterfeldt, 2015). Two particular types; outcome and confirmation bias, seem well placed to have contributed to the low levels of VIP found in this thesis. Outcome bias is described as a tendency to ignore warnings and can result in “normalising the abnormal” (New Scientist, 2015). Therefore outcome bias could be leading participants into not writing down the hazards they see, due to an underestimation of the risk presented. An example from the kitchen study could be observing fire doors propped open. These hazards can be seen so many times in the “real word” (as the lead investigator can attest to) that their occurrence can be described as common place. Due to a consequential lack of actual fires from this hazardous condition, fire doors left open may have become normalised, and not viewed as a hazard any more.

Confirmation bias has been described by Kahneman, (2011) as; experts “seeing” what they already expect to see without due regard for the evidence in front of them. So confirmation bias may also be at work here. The kitchens used were high quality with recently purchased equipment and surfaces. They were also subject to regular cleaning and in effect, these kitchens looked clean and shiny which may have pre-disposed participants to expect good standards and thereby “see” fewer hazards.

6.5.3 Hybrid Foraging Search Behaviour

This type of visual search behaviour has been experimentally shown to fundamentally limit VIP (Cain et al, 2013; Fougnie et al, 2015; Wolfe, 2013; Wolfe et al, 2016; Zhang et al, 2015). The argument that it could have limited VIP in the two thesis studies is as follows. If control group participants were searching for hazards using hybrid foraging behaviour, then marginal value theory predicts the number of hazards they observe, will reflect their previous experience of “real world” searches. In short, they will look for roughly the same number of hazards as in previous visual inspections. If their previous “real world” experience consisted of low hazard observation rates, then this same expectation could have been carried into the inspection tasks created for the two thesis studies.

There is some weak and circumstantial evidence to support this theory. During the pilot studies, the lead investigator could not help but notice a rather aimless approach to the visual inspection task by control group participants. Many of these participants did not seem to exhibit any purposeful pattern to their inspection behaviour. Rather it was seen that participants randomly wandered around the kitchen, stopping only to write something down before resuming what resembled an aimless search. Furthermore there was an impression that a number of participants began to write less and less as the 30 minute time period approached. In addition, various control group participants appeared to have ceased writing some time before the end of the inspection period. It is accepted that this is weak evidence, is subject to cognitive bias and did not include any data gathered under experimental conditions. Nevertheless, hybrid foraging search has been demonstrated experimentally as a human visual search behaviour that limits observational accuracy (Cain et al, 2013; Fougnie et al, 2015; Wolfe, 2013; Wolfe et al, 2016; Zhang et al, 2015). Therefore, if this visual search behaviour was used by control group participants, it could have detrimentally affected VIP.

6.5.4 In-Attentional Blindness

There also seems to be a compelling reason to consider the effect of in-attentional blindness (Aimola-Davies et al, 2013). This is because many participants simply did not see hazards that can only described as obvious. For example in the kitchen study, hazards such as 330ml skull and cross-bone labelled bottles of water, were not observed by just over 50% of the participants. It seems that in-attentional blindness could well be part of the explanation for why, some very noticeable hazards were simply not observed.

6.5.5 Subsequent Search Misses

To recap on this sensory perceptual cause of visual search error, attention is guided to items with similar features to targets that have already been observed (Fleck et al, 2010; Mitroff et al 2015; Wolfe et al, 2016). Therefore the characteristics of an initial target observed, somehow primes the visual inspector to observe similar targets. This could have led participants to preferentially observe hazards, similar to the ones they observed during the early stages of their trial inspections.

All the causes of visual search error listing above (sections 6.5.1-6.5.5) are presented as the most plausible speculative contenders to explain the generally low levels of VIP demonstrated in the two thesis studies. However it should be noted that given the complexity of human visual search behaviour, this listing is at best a carefully considered speculative position. It is also very probable, that a combination of these causes, with additional ones (see Table 3) and others yet discovered by visual search scholars, were most likely acting on the participants.

6.5.6 Why did SVS not Match Attentional Set Training Improvements

Although SVS did improve VIP in both the thesis studies, it did not have the spectacular success reported by Albert et al, (2017). These scholars raised VIP from 32% to over 70% using an attentional set training intervention consisting of a ten word hazard mnemonic (see section 3.2.5). SVS is a method that promises exhaustive visual search for all areas under analysis. But experimental group participants only saw about half of all observable hazards in the kitchen study, which was below expectation. However, this underperformance is only apparent when viewed against the success of Albert et al, (2017). But a closer look at this particular study can question the veracity of some of the very high levels of improvement reported (as detailed in section 3.2.5).

To begin with the Albert et al (2014) study reported a far more modest VIP increase of about 12% to a mean of 49.75% which is in keeping with the results from this thesis. In contrast their 2017 study, using the same mnemonic intervention but applied in three different pedagogies, was far more successful with an overall resultant VIP of over 70%. One pedagogy they used, that of using hazard signs on the construction site, returned a VIP of over 80%. Furthermore, these scholars used a relatively easier to master attentional set method; a ten item mnemonic when compared to the skill based method of SVS. In addition, these studies had at their disposal far more researchers to assist with training, all of which may explain the higher levels of VIP demonstrated, relative to the two thesis study interventions.

6.5.7 Possible Reasons for SVS Underperforming

Even though the intervention did significantly improve VIP using only a thirty minute training session, it remains that participants using the SVS method could have underperformed. There are a number of plausible explanations for this possibility. Participants may in some instances not have understood the method or did not use it as intended. Alternative eye scanning patterns may also have improved results. However on long reflection, there seems to be two possible and plausible but as yet speculative explanations that could have combined to limit SVS results and explain this underperformance.

The first is that the results could be simply reflecting a lack of visual inspection practice in using SVS. VIP for control participants was found to be circa 32% and 38% in the two thesis studies. It therefore seems intuitively difficult to see how a relatively short 30 minutes training session in the skills based SVS method, would raise VIP scores to the levels achieved by Albert et al. (2017). Furthermore, SVS represented a novel inspection skill to be learned in 30 minutes by participants. On learning the SVS method, these same participants were immediately taken to the observational task and directed to use the method for the first time in their careers. With this in mind, it can be argued that SVS did well in achieving a circa 17% & 25% increase in VIP from the two kitchen and aviation studies.

The other reason could be due to SVS users rushing their observational tasks or running out of time. As exemplified by Cain et al, (2013), any type of time pressure in the visual search process increases error. The argument for rushing is plausible and supported by two further pieces of evidence. Firstly it took the lead investigator about 20 minutes to conduct a visual inspection of the two pilot trial kitchens, and to confirm the number of hazards prior to each of the 12 kitchen trials. Given the experience of the lead investigator’s 30 years of visual inspection conduct, it seems plausible if not probable, that experimental participants needed some of the available inspection time; to master the SVS skill and come up to speed. If this occurred, these participants may have rushed the visual inspection or ran out of time. The thesis data also shows that experimental participants wrote down more hazards and thereby had less time to visually inspect when compared to their control colleagues (see sections 5.1.6

& 5.4.6). So by any metric, the rate at which experimental group participants were working, must have been at a faster rate than their control colleagues. This may have contributed to increases in visual search error, which may in turn have limited VIP amongst SVS users.

6.6 Should SVS be Promoted within the EHS Community

The next section of this discussion explores whether or not SVS has a potential role to play in relation to EHS practice. This section begins by listing the disadvantages of SVS. It will then ask whether the thesis results can be generalised and if so, why are there so few incidents from all these un-observed hazards in workplaces. It concludes by considering possible effects from any un-observed workplace hazards and what the EHS community may think of this new method.

6.6.1 An Increase in Time to Conduct Competent Visual Inspections

The main methodological drawback from SVS seems to be that it will take more time to conduct appropriate visual inspections, thereby increasing professional costs. The kitchen study data points to upper limit of about one hour needed for every $100m^2$ of hazard rich, workplace floor area. This estimation is arrived at as follows. The mean floor areas of the kitchens were $95m^2$, so rounding to $100m^2$ is reasonable. At the lower end of the time scale required, the lead investigator visually inspected the kitchens before each trial, and wrote down his notes in about 20 minutes. In the “real world” these notes would need to be read over, to ensure clarity of writing and to add any conclusions thoughts and reflections. So a lower period or 30 minutes for an experienced SVS user would seem a reasonable amount of time to allow a $100m^2$ workplace such as a kitchen, to be adequately visually inspected.

The numbers of hazards written by experimental participants was a mean 48.86 and 52.15 for the kitchen and aviation studies respectively. So close on fifty hazards can be written down within the half hour period. However the positively skewed platykurtic distribution of these scores means that just under half of participants will need more than thirty minutes. So a practical solution to ensure sufficient time for those visual inspectors who are to the left of the median in the VIP distribution curve, would be to increase this half hour period to the next pragmatic working time period of one hour.

6.6.2 Importance of Hazards not Differentiated

A further disadvantage of SVS is that using a search strategy to fixate on observable hazards, does not differentiate them in terms of importance. This means that a low risk hazard is as likely to be observed as a high risk hazard. This is not in itself a bad thing except in the situation whereby time is short. In such instances high risk hazards are as likely to be observed as a proportion of all hazards, and not with the priority their importance requires. However it is difficult to see how any observational method could prioritise hazard importance.

6.6.3 Generalising to the EHS Community

Before deciding whether or not SVS can be promoted within the wider EHS community, a consideration is required to see if the two thesis studies can generalise. The main reasons for suggesting generalisability are; the robust experimental design, statistical power, ecological validity, and replication by the aviation study.

However there is a strong argument to mitigate against generalisability namely population homogeneity. As stated in section 6.2.1, thesis participants were not fully representative of the wider EHS community which consists of a very diverse range of disciplines. With the kitchen study, it can also be argued that using undergraduates is a flawed strategy as they are not representative of the wider full time working population. It can further be argued that the very fact that these participants collaborated, shows that they self selected themselves and are thereby showing a level of motivation higher than their non-participating colleagues.

However there is an overriding argument that could sway a conclusion towards generalising to the wider EHS community. It is accepted that the sample was not representative, but this is a statistical conclusion based on the fact that a random sample was not recruited and geographical diversity was not achieved. However it is argued, that the percentage density functions showing VIP (Figures 11, 12, 15 & 16) are reflective of a human characteristic. Furthermore the thesis results have been supported by studies from North America, the UK and Ireland all of which, evidence our collective human flaws when it comes to accurate visual inspection (Albert et al, 2014 & 2017; Drury & Watson 2002; Moore et al, 2001; Hollis & Bright, 1999; Hrymak et al, 2015). Therefore, it is difficult to refute the evidence from the two thesis studies due to uncertainty over sample representativeness. It does not seem particularly convincing to argue that a basic human characteristic, namely the way humans visually search for defects, targets and hazards will be different for EHS professionals based on their jurisdictional location.

It is however, far more convincing to say that EHS professionals will reflect what is scientifically evidenced about the way humans visually search for defects, targets and hazards. On that basis, it seems that there is more weight of evidence to generalise the findings, than there is to ignore them until better sampling is achieved. Put another way; should the two thesis study findings be precluded from influencing the wider EHS community and allow the unchallenged and continuing practice of observable workplace hazards being missed. The intuitive answer to this seems to be no. Therefore, there seems to be a moral argument at the very least, to suggest the two thesis study findings are brought to the attention of the EHS community, even if it just to begin a debate as to what constitutes competent VIP.

6.6.4 Why So Few Accidents if Visual Inspection is Below Expectation

There is a further, perfectly logical argument, to negate generalisation from this thesis and the literature. If EHS practice is characterised by VIP rates in the thirties region percentage wise, why are there not far more accidents, illnesses and economic loss incidents from all these unobserved hazards. In short, if there are so many hazards being missed, why are we not seeing far more detrimental effects?.

In answer to this argument, there are perfectly good reasons to explain how current visual inspection custom and practice, and the number of accidents, ill health and economic loss incidents can co-exist. But at present, there is only anecdotal evidence and circumstantial literature to account for this reasonably made point. One argument for this co-existence, is that two professions, already use consistency in their visual search behaviour; Chartered Surveyors conducting building surveys and medical practitioners conducting physical examinations. These two professional groups will conduct their visual inspections using a consistent search strategy. This group of professionals do not have the kind of publicity that is associated with a VIP rate in the thirties range percentage wise. Therefore there is circumstantial evidence for improved VIP with the use of a SVS type strategy. Furthermore these professional searchers have the advantage of a well understood and described attentional set. In effect, they have a limited number of defects, hazards and targets to look for. For example Chartered Surveyors have a set number of building defects such as damp, wood rot, structural movement etc. In contrast EHS professionals have a potential hazard listing that is characterised by some 15,000 terms and synonyms (International Labour Organisation, 2017).

Another argument is that the number of accidents, ill health and financial loss incidents, actually does reflect the non-observation of observable workplace hazards. Consider that in Ireland in 2015 there were 56 work related fatalities. There were also 18,796 persons injured at work to the extent that they were absent for more than 4 days (HSA 2017). In that same year, there were 6,707 personal injury claims (Health and Safety Review, 2018) So why did these fatalities and injuries occur if risk assessments were conducted. There are enough adverse incidents here, to propose that a proportion of these could have occurred due to workplace hazards being overlooked.

There is also the redundancy argument that states, a certain amount of safety is built into systems, especially mechanical components (Drury & Watson, 2002). This redundancy argument is especially applicable to the aviation study. Here informal conversations between the two senior aviation engineers and the lead investigator revealed that the high level of safety built into aviation components can in many cases, allow for their failure without automatically compromising flight air-worthiness.

A further argument is given by Reason, (1997). Hazards have a potential, and not a set probability to cause accidents, ill health and loss incidents. Luck plays its part as companies with good safety management can have accidents, and those with poor safety management may not have any (Reason, 1997). So the observation, or not, of a workplace hazard does not necessarily translate into a dichotomous and predictable; “the accident will or wont happen” scenario.

Another argument centres on one of the oldest theories in risk management, Heinrich’s accident triangle which dates back to 1931. Although this safety pioneer’s figures are not now accepted, his theory still centres on the relative frequencies of minor, major and fatal accidents (Marshal, et al 2018). Minor accidents are at the bottom of this pyramid due to their large number, and fatalities at the top due to their infrequent occurrence (Marshal, et al 2018). The argument here is that workplace fatalities are thankfully far less numerous than non-fatal injuries. As detailed above for 2015, there were 56 work related fatalities and 18,796 persons injured at work (HSA 2017). SVS does not of itself, distinguish as to the potential of the observed hazard to cause minor or major accidents. Therefore SVS will clearly result in proportionally far more of the less serious potential hazards being observed, simply because of their much higher relative frequency in the workplace.

A final axiomatic argument is to consider the examples of visual inspection failure that are detailed in sections 1.2.3 & 1.2.4. Here tragic examples of preventable fatalities together with workplace injuries and substantial economic loss are detailed. In summary, not observing workplace hazards during workplace inspections has been demonstrated to cause accidents, ill health and financial loss. Therefore and in summary, there are good reasons to suspect that poor VIP and current rates for accidents, ill health and financial loss incidents, can co-exist.

So a principal concern from the thesis findings remains; the relatively low level of visual search performance, together with individual variability exhibited by the participants. Standing in the thirties region percentage wise, VIP is far too low when measured by any professional yardstick. Even if this figure is taken to region 50-60% as exemplified by participants using SVS the results are still cause for concern. The implication when extrapolating to the wider EHS community does not make easy reading. Simply put, EHS professionals are putting themselves in a position where they are routinely not observing some two thirds of all workplace hazards that professional standards require. The consequences of routinely missing observable workplace hazards are clear. The possibility of accidents, ill health and loss incidents occurring must be higher than if these hazards are observed and thereby acted upon.

The literature on accident causation has consensus and frequently relies on the work of Reason, (1997). This scholar published the seminal “swiss cheese” accident causation model. This model describes un-identified hazards as being in a latent state, but will cause harm when a number of conditions align and allow the hazard to transition into an adverse event. Simply put, un-identified hazards or in the context of this thesis, un-observed hazards are accidents waiting to happen. So the point being made here is that latent hazards have consensus in the EHS community as having potentially negative consequences. Therefore, they are in need of identification and remediation. In short, un-observed workplace hazards are a problem and any method that can improve hazard identification should be welcomed.

6.6.5 Possible Responses from the EHS Community

A number of qualitative comments have been recorded by the lead investigator, when informally discussing the thesis results with academic and professional colleagues. If these comments are anything to go by, then the possibility of a lukewarm response from this community exists. If so, it is clear that any improvement in the conduct of visual inspections will require academic and not professional leadership. These comments are as follows;

“there’s no way I’ll spend more time doing inspections, clients just wont pay for it”

“ I can get five risk assessment done in a day, you’re telling me I need to cut that down to one, there’s no chance of that”

“ I don’t need your method, I can smell hazards a mile off”

“your approach is caveman, auditing and checking paperwork is the modern way”

CHAPTER 7 CONCLUSIONS & RECOMMENDATIONS

7.01 Introduction

This chapter will begin by reiterating the core conclusions that can be taken directly from the thesis. This will be followed by those conclusions that are more speculative in nature but can be drawn from the thesis together with literature. A section on recommendations for visual inspection practice and inquiry will then be detailed and the chapter will conclude by outlining potential roles for the systematic visual search method.

7.1 Direct Conclusions from the Two Thesis Studies

All the following are core conclusions that can be supported by reliable and valid data from the two thesis studies. To recap regarding abbreviations, SVS refers to systematic visual search, VIP to visual inspection performance and EHS to environmental health and safety.

1. The findings from the two studies in this thesis demonstrate that the use of systematic SVS leads to significantly better VIP than current custom and practice.
2. In overall terms, VIP from all participants was below expectation, and is therefore of concern from an EHS professional practice perspective.
3. SVS demonstrated two efficiency gains over current visual inspection custom and practice, users observed more hazards in less time and users wrote down more hazards.
4. Percentage density functions for visual inspection performance and hazards written down, were platykurtic shaped distributions (negative kurtosis or sharpness of peak, with light tails).
5. There were no hazard salience characteristics found that predicted observation. How much the hazard stood out from its surroundings did not seem to affect how often it was observed.
6. The SVS method can be learned in 30 minutes. As such it represents a cost efficient method for improving visual inspection practice.

7.2 Conclusions Supported by the Wider Literature

There are a number of more speculatively grounded conclusions that can be presented by combining the two studies in this thesis, together with the literature detailed in Chapters 2 and 3.

1. The utility of SVS seems to occur in conjunction with efficiencies in cognitive resources.
2. SVS seems to be utilising bottom up stimuli processing to guide visual attention and deployment, rather than any top down stimuli or scene guidance (only looking in those areas where you normally expect to see hazards).
3. By using bottom up stimuli processing, SVS would appear to be less reliant on environmental conditions. The method thereby seems to possess potentially universal application.
4. By promoting a higher level of hazard observation, SVS may have exhibited the ability to lessen outcome and confirmation bias effects. If more hazards are observed then this could challenge the visual inspector’s decision analysis process to better accept visual evidence, and counter such bias.

7.3 Further Research

Having reflected on the methodology and the subsequent data analysis, future avenues of inquiry to improve visual inspection practice can be suggested.

7.3.1 Controlling Further Variables within the Experimental Design

A number of variables could be further controlled using the thesis experimental design. For example, if time taken to inspect, feedback, the eye scan pattern used, and practice are controlled as independent variables then additional data would be generated as to best practice in visual inspection. A further useful intervention would be a comparison between checklists, using questions and on site document analysis against SVS. In this way empirical evidence as to the utility of all these inspection methods can be evaluated. This may also lead to a ranking of utility for inspection methods.

7.3.2 Using Qualitative Data

Qualitative data is largely absent from this thesis. Therefore there is little evidence in this thesis, of just how beneficial EHS professionals would actually find SVS. There can be little doubt that qualitative data following an interpretivist phenomenological narrative epistemology, would be very useful for future visual inspection inquiry. In short, the social reality of EHS professionals actually using SVS in their day to day duties, will greatly strengthen research in this field. Following on from this, it will be advantageous to interview experimental group participants to ascertain their experience of SVS behaviour during the inspection tasks. This qualitative methodology would reveal valuable data as to how closely participants stuck to the method as well as what advantages and disadvantages they encountered with its use.

A closer analysis of what participants wrote down in their scripts will deliver further interesting data. One additional area of inquiry could have investigated if control group participants were exhibiting foraging behaviour as predicted by Cain et al, (2013); Fougnie et al, (2015); Wolfe, (2013); Wolfe et al, (2016;) Zhang et al, (2015). Another area of inquiry from the scripts, would have been to see exactly what were participants writing about if observed written hazards are removed from scripts.

7.3.3 A Comparison Between Systematic Visual Search & Safety Auditing

Safety auditing, is a very widely used risk assessment method. As with risk assessment conduct, precise meaning is largely absent. However, a methodological commonality of auditing, involving a tripartite paradigm of “look ask & read” has been described by Hrymak et al, (2015). In auditing practice, the use of questioning workplace employees and reading on site documentation is widespread and used in an attempt to identify hazards and evaluate safety (Blewett & O’Keefe, 2011). SVS is also a method of assessing safety, as it directly allows for the observation of workplace hazards. So it would be interesting to compare the utility of auditing with SVS. An experimental design could include further workplaces chosen for inquiry for example, construction sites, mechanical workshops or manufacturing facilities.

If for example a construction site is used, then researchers posing as EHS professionals could invite auditors to use their customary practice of asking questions, reading documents and looking around the site. If the construction site is set up as per the thesis experimental design then it will contain a known number of hazards. The audit can be then evaluated in comparison to SVS being applied to the same site.

7.3.4 Investigating Scene Deployment Effects

One speculative theory to explain control participant visual inspection behaviour is the effect of scene deployment. To recap from section 2.1.1 this type of guided visual attentional deployment involves tending to looking most closely, where you expect to find your objects of interest. As scholars explain; asked to look for fish, most people will look for a water body and not at the sky (Wolfe & Horowitz, 2004). If visual inspection behaviour is influenced by scene deployment, then this can be confirmed by the precise placement and distribution of hazards. Therefore, using an engineering workshop or a manufacturing facility as an example, hazards could be planted in areas where they would and would not be expected. In this way any such scene deployment effects, can be evaluated.

7.3.5 The Need to Evaluate Which Training Curricula to Use for VIP

All the training interventions including the use of feedback, from the visual search literature report positive effects. See for example Albert et al, (2014 & 2017); Drury & Watson, (2002); Nickles et al. (2003); Rao et al. (2006); Schwaninger, (2005); Wales et al, (2009). These scholars all report improvements in VIP from training. So a good question here is exactly what is the best syllabus and educational pedagogy to maximise VIP. At present, this question is difficult to answer with any confidence. Attentional set training would seem the least amenable to EHS professionals as it can be assumed that they already possess a high level of understanding of workplace hazards. Therefore, there is a clear need to evaluate the effect of different curricula and pedagogy on VIP in order to recommend the most appropriate training. In this regard the experimental design as used in this thesis, could potentially serve as a test bed for differing educational approaches and curricula.

7.4 Improving Visual Inspection using SVS

This section looks at what SVS may offer in practical terms, in order to improve current custom and practice for visual inspection.

7.4.1 Tell Visual Inspectors there are 30 Hazards to be Found

To recap from Chapter 2 (Theoretical Underpinnings) scholars have demonstrated that if hybrid foraging search behaviour is used recent VIP can influence subsequent search performance. (Cain et al, 2013; Fougnie et al, 2015; Wolfe, 2013; Wolfe et al, 2016; Zhang et al, 2015). Theoretically, if EHS professionals can be anchored to a higher number of hazards then SVS may offer a way of observing this higher number. So one pragmatic way to improve current visual inspection custom and practice is to suggest that EHS professionals should use SVS to find 30 or so hazards, in any workplaces they visually inspect. This number of 30 has been chosen from the kitchen study trials. A higher number would intuitively, seem excessive. A lower number would be problematical given that 30 hazards can very easily be present as seen in each of the thesis trials conducted. This anchoring around the number 30 should in theory be self-reinforcing (Kahneman, 2011). The idea here being that if the visual inspector does observe 30 or so hazards per inspection, then such a high VIP recency rate should translate into subsequently higher rates of VIP from hybrid foraging searchers.

7.4.2 Educational Awards to Recognise Visual Inspection Competency

The precise proceduralisation, of visual inspection conduct that SVS brings, together with the use of the simulated inspection experimental design used in this thesis, provides EHS educators with a practical method of assessing VIP. For example candidates for EHS educational awards could be required to conduct actual workplace inspections using the SVS method under normal educational examination conditions. In this way assessment criteria can easily be set to ensure professional competence together with practice and feedback requirements into curricula.

7.4.3 Lessening the Ubiquity of Checklists for Visual Inspections

Checklists are ubiquitous, easy to use and rapid (Clift et al 2011; Neathey et al, 2006). However they remain visual search methods and as such, will suffer from the same human cognitive limitations outlined in this study and summarised in Table 3. In addition, checklists can present further problems from increased time requirements and a selection dilemma in terms of which checklist to use. To explain; many checklists require multiple searches of the same element. A common example would be a search of the floor for trip hazards, and then searching the floor again for cleanliness. Hence the floor will be visually searched twice if not more times depending on what the checklist requires. These multiple visual searches of the same area will increase inspection time.

Another limitation of checklists is intentionally or unintentionally omitting hazards. Some hazards will inevitably have to be omitted, due to the sheer scale of potential workplace hazards. Various commentators including Aven, (2011); Clift et al, (2011); Neathey et al. (2006) point out that checklists only cover the hazards written down. If they are not written down, then it is up to the EHS professional to note any omissions and act accordingly which introduces further uncertainty over the visual inspection. But covering all hazards in a checklist is impossible to do in practice considering that the International Labour Organisation, (2017) lists some 15,000 terms and synonyms used to describe hazards. In short omitting hazards from checklists is a certainty. Various scholars voice their support for checklists usually due to their simplistic utility (see Gawande, 2009; Neathey et al. 2006) However there is a counter view stating that checklists are only useful at the end of any risk assessment process to ensure no hazards are overlooked (Aven, 2011). In short, lessening the reliance on checklists by offering SVS as an alternative would potentially improve visual inspection conduct. A further strand to this inquiry is to evaluate this advice from Aven, (2011). An experiment to evaluate the effect of SVS on for example, a mechanical workshop compared to the same workshop but with a checklist directed visual inspection, could isolate the efficacy of the checklist used.

7.4.4 Standardising Visual Inspection Conduct

For professionals working within the EHS community, there are transparency validity and reliability advantages to the use of SVS. This method also fits in well with the requirement of Le Coze, (2005) who calls for risk assessment methods that are; not so complex that they become far too time consuming and expensive, whilst they are not so simplistic as to render findings superficial, generic and failing to identify workplace hazards.

Currently visual inspection custom and practice is idiosyncratic and does not have consensus professional proceduralisation or oversight (Hrymak et al, 2015). In effect, no one to any real extent checks the checkers, audits the auditors or risk assesses the risk assessors. Furthermore professional and official guidance on how to conduct workplace inspections remains vague to say the least. SVS therefore offers the possibility of very specifically and strictly proceduralising visual inspection conduct. As such, it offers the EHS community a potentially standardised method for the conduct of the vast number of workplace inspections and subsequent risk assessments carried out globally.

In support of reducing ambiguity and addressing the bewildering array of phrases used to describe workplace visits for safety reasons (see section 1.1) a consensus definition on visual inspection would seem to be advantageous. In short, rather than saying a workplace risk assessment was conducted, EHS professionals could be provided with a far more precise meaning to their visit if they could state; a risk assessment was conducted which included a visual inspection. To this end a precise definition of visual inspection could therefore be;

"An exhaustive visual search for multiple instances of defined multiple hazards, for a precisely specified area or object"

7.4.5 Using Systematic Visual Search for Multi-Screen Monitoring

This is an interesting recent development that could provide SVS with further potential EHS applications namely, multiple object tracking. A study by Wu & Wolfe, (2016), found that that observers can only see two or three screens at a time with any accuracy. This has consequences for control room management, for example watching multiple security cctv (closed circuit television) screens in monitoring centres. Control rooms are common place for multi-site monitoring organisations and in safety critical industries. This thesis has demonstrated an visual search efficiency effect from SVS, when observing hazards. Therefore there is potential to apply SVS to multiple object tracking in control rooms. SVS could be applied here by using its selection of objects prior to eye scanning pattern proceduralisation. Take for example a bank of cctv monitors mounted on a control room wall in a fire alarm monitoring company. The "reading a book" eye scanning pattern will direct the visual inspector to select the top left monitor on the wall. The same eye scan pattern can be used to visually search this screen before the next monitor is selected, and so on until all monitors are searched.

7.4.6 Extending SVS to Evaluate Safety Critical Human Behaviour

SVS has been shown to be applicable to workplaces and objects. In theory there seems to be no reason why SVS would not be applicable to observing human behaviour for example manual handling behaviour of baggage handlers. If SVS is extended to evaluate its capability to observe human behaviour, it can extend its potential utility as well as providing accurate safety related observations. In effect, it could have an important role in job safety analysis, where employee observation is required. The use of observation in job safety analysis is established and a listing is provided by Tomaschek et al, (2018). These scholars speculate that observation can be extended to certain areas of psychosocial risk assessment from in-appropriate workload demands leading to stress. These scholars note that whilst workload is typically difficult to observe, components that are factors affecting workload are observable and include; task conflicts, work interruptions and multitasking (Tomaschek et al, 2018). Hence these observable work practices can potentially fall within the remit of SVS, which has demonstrated its ability to deliver improved observational accuracy.

7.5 A Succinct and Finalised listing of Recommendations

Taking the literature and findings from this thesis, a succinct listing recommendations will now be presented under the headings of; further research, generalising SVS, defining visual inspection competency and scenarios where SVS should be used.

7.5.1 Further Visual Inspection Research to Maximise SVS Potential

As section 7.3 above describes there are a number of research agendas that could be pursued with the ultimate aim of maximising visual inspection performance when using the systematic visual search method. This research should be conducted with a real world ecologically valid emphasis, using interdisciplinary researchers from the natural and social sciences.

7.5.2 Generalising to the wider EHS Community

There is now seems to be enough evidence from the literature and this thesis studies for the EHS community to consider field trialling SVS. As discussed in section 6.6.5, this community may well be sceptical about SVS. If this is the case, it suggests that an academic led research programme will be required to evaluate and disseminate SVS use in professional practice. At the very least a debate in the professional Environmental Health and Safety community needs to begin on the merits of current visual inspection practice and the over-reliance on checklist methods.

7.5.3 Defining Competence for the Visual Inspection of Workplace

The literature in section 2.5 illustrates how the inherent ambiguity for visual inspection, risk assessment and safety auditing conduct, is responsible for idiosyncratic workplace inspections. The proposed definition for visual inspection from section 7.4.4 would be an important and beneficial first step in standardising visual inspection conduct and providing a competency criteria for EHS professionals. To recap, this proposed definition is; “An exhaustive visual search for multiple instances of defined multiple hazards, for a precisely specified area or object”.

7.5.4 Recommended Scenarios for SVS use

As discussed in section 6.1, SVS has the potential to offer improved observational accuracy for EHS professionals. As such, SVS can be recommended for situations where workplace visual inspections are conducted for risk assessment and safety auditing purposes. The following considered listing, illustrates specific situations in which EHS Professionals could use SVS, in order to increase their visual inspection performance. These situations include;

- Inspections of health care facilities by the Irish statutory body HIQA (Health Information and Quality Authority.
- Infection Control and prevention staff in Health Care settings when inspecting premises for hygiene purposes.
- Food service staff such as Chefs and managers who are tasked with daily inspections of their working premises for safety and hygiene reasons.
- EHS professionals working with organisations that present high incidences of personal injury claims such as Local Authorities, the Health Service Executive, and educational providers.
- Visual inspection conduct for those statutory bodies that are required by law to visit workplaces such as the Health and Safety Authority inspectors, Environmental Health Officers, Local Authority Fire Officers, Radiological Protection Inspectors, and Environmental Protection Agency Inspectors.
- Meat inspectors and veterinarians assessing carcasses from a fitness for human consumption perspective.
- Defence forces personnel looking for improvised explosive devices in conflict zones, and search and rescue crews using aircraft to look for marine casualties.
- An Garda Siochana investigators conducting crime scene and forensic searches.

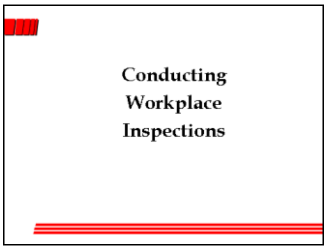
7.6 In Final Conclusion

The aim of this thesis was to develop and trial Systematic Visual Search and there is now empirical evidence for its beneficial use. Therefore in terms of reducing workplace accidents, ill health and loss incidents, systematic visual search has a role to play. The promotion of systematic visual search within the wider EHS community is now required. In this regard, the lead investigator will continue to disseminate the findings of this thesis, conduct further visual inspection research and promote this behavioural visual search algorithm amongst professionals. In addition, this dissemination and promotion can now continue with the added confidence that stems from reliable and valid data, supporting the beneficial effects of systematic visual search.

APPENDICES

Appendix 1; Recapping Lecture Slides; Used by All Participants

The power-point slides used for the recapping session on kitchen food and safety hazards were as follows;



Conducting
Workplace
Inspections

.....

...

......

......

......

...

.....

.....

.....

.....

(c) pcspestcontrol.co.uk

.....


NILFISK

.....

.....


FIRST
AID

.....

.....

.....

.....

.....

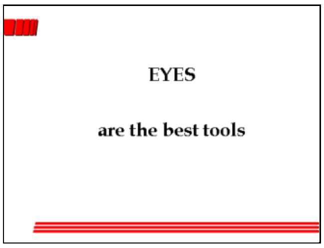
Appendix 2; The Syllabus for SVS

The following syllabus was delivered during the 30 minutes available to experimental group participants

- Visual search is a common every day occurrence
- Successful search requires fixation by the eye on the object of interest
- Visual search can be random or systematic
- Successful search of all relevant targets is not easy
- Successful search requires a consistent or systematic eye scanning strategy
- Fixation using the foveal region will allow for maximum acuity
- The eye scanning pattern to use is the “reading a book” pattern
- It is important that the eyes need to see or “paint” all surfaces in the room
- Consistency ensures no areas will be visually missed or searched repeatedly
- Consistency is achieved by order of search; ceiling, wall, floor, storage & equipment
- Consistency means eye scanning only the element selected and to completion
- Move push pull walk crawl, go on bended knee if direct line of sight not achieved

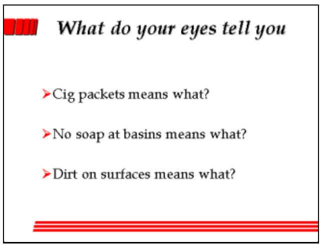
Appendix 3; Experimental Group Lecture Slides

As the participants were in the same lecture hall, the power-point slides used for this group were exactly the same re-capping slides. The following additional slides were used for the SVS pedagogy;



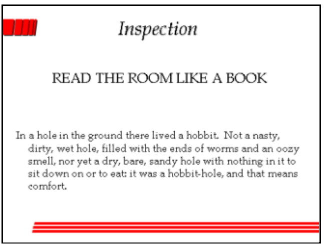
EYES

are the best tools



What do your eyes tell you

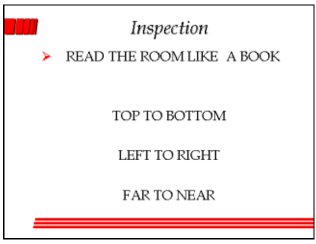
- Cig packets means what?
- No soap at basins means what?
- Dirt on surfaces means what?



Inspection

READ THE ROOM LIKE A BOOK

In a hole in the ground there lived a hobbit. Not a nasty, dirty, wet hole, filled with the ends of worms and an oozy smell, nor yet a dry, bare, sandy hole with nothing in it to sit down on or to eat: it was a hobbit-hole, and that means comfort.



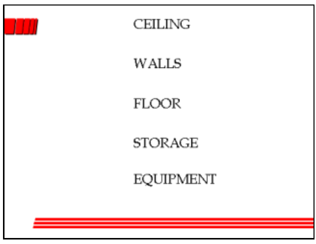
Inspection

- READ THE ROOM LIKE A BOOK

TOP TO BOTTOM

LEFT TO RIGHT

FAR TO NEAR



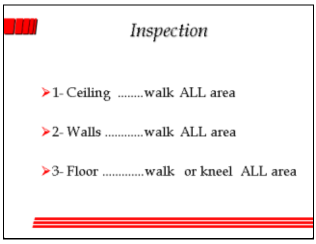
CEILING

WALLS

FLOOR

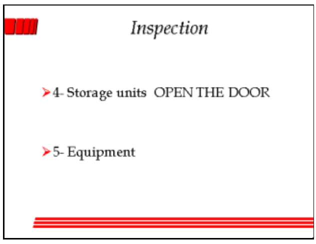
STORAGE

EQUIPMENT



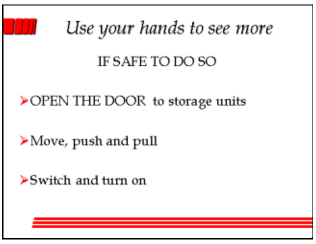
Inspection

- 1- Ceiling ....walk ALL area
- 2- Walls ....walk ALL area
- 3- Floor ....walk or kneel ALL area



Inspection

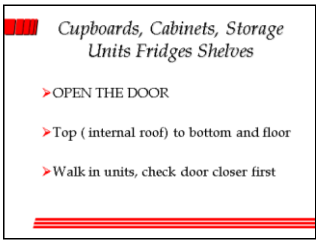
- 4- Storage units OPEN THE DOOR
- 5- Equipment



Use your hands to see more

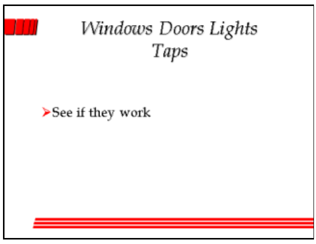
IF SAFE TO DO SO

- OPEN THE DOOR to storage units
- Move, push and pull
- Switch and turn on



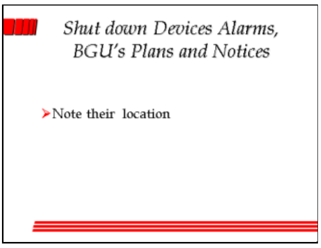
Cupboards, Cabinets, Storage Units Fridges Shelves

- OPEN THE DOOR
- Top (internal roof) to bottom and floor
- Walk in units, check door closer first



Windows Doors Lights Taps

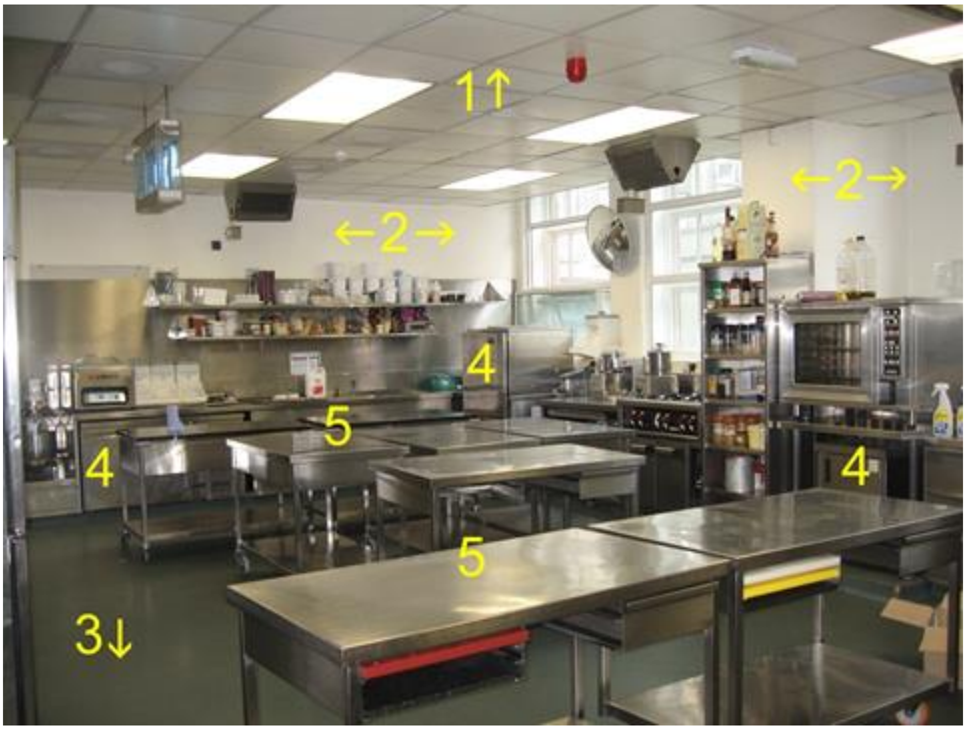
- See if they work



Shut down Devices Alarms, BGU's Plans and Notices

- Note their location

Appendix 4; Documentation for the Experimental Group



Gender; Job title ; Number of years working.

Inspection

Top to Bottom, Left to Right Far to Near

1 = Ceiling

2 = Walls

3 = Floor

4 = Storage; OPEN THE DOOR Fridges Freezers Cupboards

5 = Equipment

Open, Move, Push, Pull, Turn On, Switch On

Page 2 Ceiling, Page 3 North Wall, Page 4 East Wall, Page 5 South Wall, Page 6 West Wall, Page 7 Ceiling floor, Page 8 Storage Page 9 Equipment

Appendix 5; A Listing of the Kitchen Study Hazards

Hazard	**Location**	**Risk Category**	**Risk Rating**	**Field of View**	**Provenance**
Covered or interfered smoke detector	Ceiling	OSH	High	Open	Planted
Obfuscated wall surfaces not readily cleansable	Walls	Food	Medium	Manipulate	Existing
Wall or shelf surfaces not readily cleansable	Walls	Food	Medium	Open	Existing
Mouse traps under equipment	Floor	Food	High	Manipulate	Planted
Moved ceiling tiles	Ceiling	Food	Medium	Open	Planted
Ceiling tiles in disrepair	Ceiling	Food	Medium	Open	Existing
N= 10 tea candles	Shelves	OSH	Medium	Open	Planted
Flammable liquid in 330ml bottle	Shelves	OSH	Medium	Open	Planted
Flammable aerosol can 250 ml	Shelves	OSH	Medium	Open	Planted
Toxic liquid in 330ml bottle	Shelves	OSH	High	Open	Planted
Window ledges left unhygienic	Walls	Food	Medium	Open	Existing
Bare wires	Walls	OSH	High	Manipulate	Planted
Cigarettes on shelves	Shelves	Food	High	Open	Planted
Cigarettes in drawers	Storage	Food	High	Manipulate	Planted

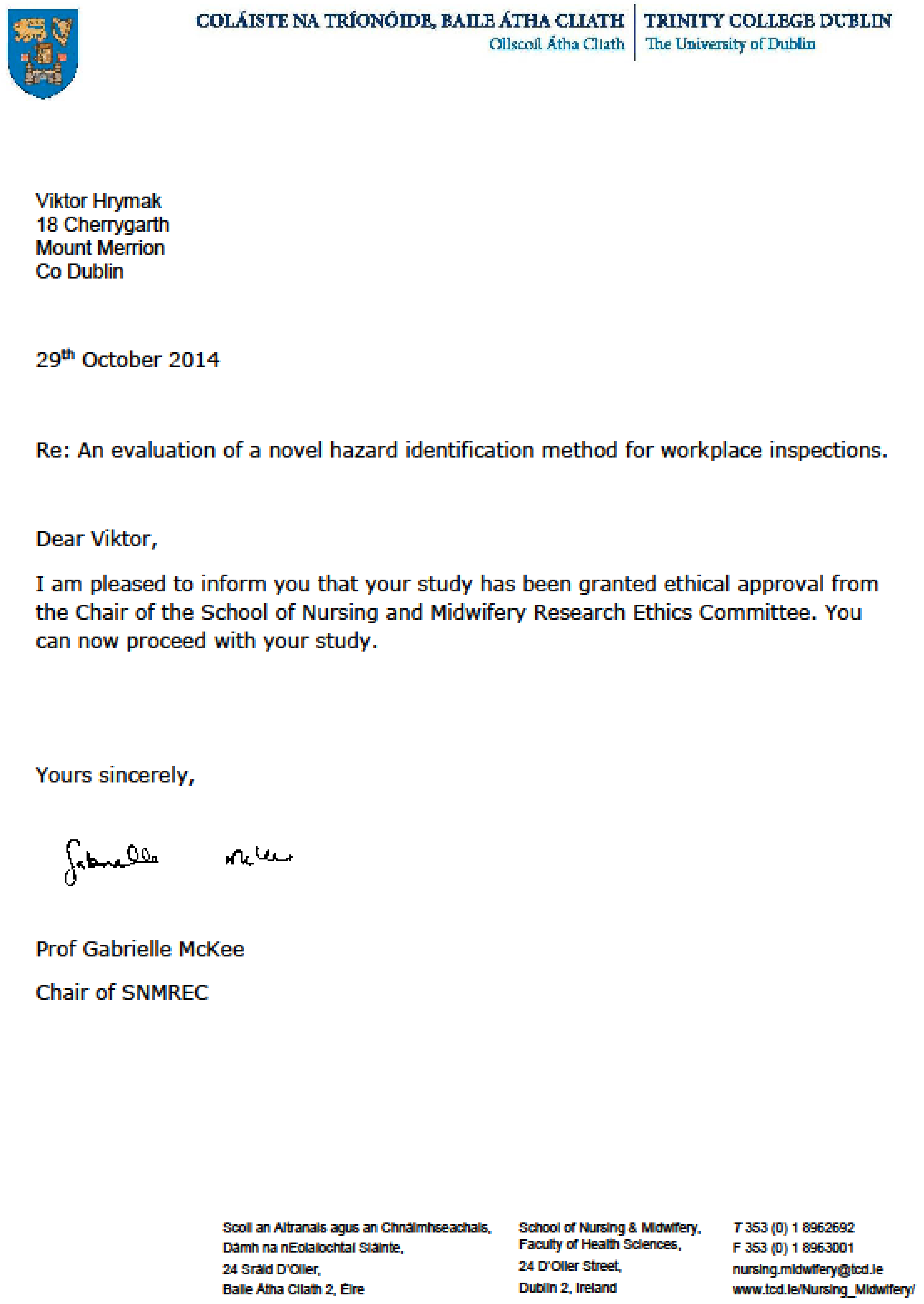
High level storage	Shelves	OSH	Medium	Open	Planted
Open food	Shelves	Food	High	Open	Planted
Overloaded sockets	Shelves	OSH	Medium	Open	Planted
Wedged fire door	Walls	OSH	High	Open	Planted
Moved fire extinguisher	Walls	OSH	High	Open	Planted
Floor damage	Floor	Food	Medium	Open	Existing
Food on floor	Floor	Food	Medium	Open	Planted
Floor left unhygienic	Floor	Food	Medium	Open	Existing
Cross contamination in fridge or freezers	Storage	Food	High	Manipulate	Planted
General hygiene	All areas	Food	Low	Open	Existing

Appendix 6; A Listing of the Aviation Study Hazards

No.	Description of Hazard	Hazard or Plant	Importance
1	Propeller blades scraped	Hazard 1	Low
2	Propeller spinner screw loose	Hazard 2	Low
3	Propeller spinner screw removed	Plant 1	Low
4	Starboard engine cloth on intake	Plant 2	High
5	Port engine lock wire broken	Plant 3	High
6	Port engine screwdriver on cylinder	Plant 4	High
7	Aerial sealer removed	Hazard 3	Low
8	Aerial screw removed	Plant 5	Low
9	Starboard wing tip tank chain removed	Plant 6	Low
10	Starboard wing filler cap cover fastener bent	Hazard 4	Low
11	Starboard wing fuel cap earth wire removed	Plant 7	Low
12	Port wing underside inspection panel removed	Hazard 5	Low
13	Port wing tip fuel tank left open	Plant 8	Low
14	Port wing corrosion	Hazard 6	Low
15	Port wing tip tank corrosion	Hazard 7	Low
16	Fuselage elevator trim tab screw removed	Plant 9	Low
17	Port elevator trim tab locking wire removed	Plant 10	Low
18	Landing gear earth wire not grounded	Hazard 8	Low
19	Port main landing gear split pin removed	Hazard 9	High
20	Port main landing gear split pin left unsafe	Hazard 10	Low

21	Starboard main landing gear flat tyre	Hazard 11	High
22	Starboard main landing gear loose bolt	Hazard 12	High
23	Cockpit pen left on windscreen	Plant 11	High
24	Cockpit circuit breakers left tripped	Plant 12	Low
25	Cockpit landing gear jack cover not secure	Hazard 13	High
26	Cockpit screwdriver left in side pocket	Plant 13	High
27	Cockpit mobile phone left in side pocket	Plant 14	High
28	Cockpit coins left on floor	Plant 15	High
29	Stall warning screw loose	Hazard 14	Low

Appendix 7; Ethical Approval



COLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH
Ollscoil Átha Cliath
TRINITY COLLEGE DUBLIN
The University of Dublin

Viktor Hrymak
18 Cherrygarth
Mount Merrion
Co Dublin

29th October 2014

Re: An evaluation of a novel hazard identification method for workplace inspections.

Dear Viktor,

I am pleased to inform you that your study has been granted ethical approval from the Chair of the School of Nursing and Midwifery Research Ethics Committee. You can now proceed with your study.

Yours sincerely,

Prof Gabrielle McKee

Chair of SNMREC

Scoil an Altranais agus an Chnáimhseachais,
Dámh na nEolaíochtaí Sláinte,
24 Sráid D'Olier,
Baile Átha Cliath 2, Éire

School of Nursing & Midwifery,
Faculty of Health Sciences,
24 D'Olier Street,
Dublin 2, Ireland

T 353 (0) 1 8962692
F 353 (0) 1 8963001
nursing.midwifery@tcd.ie
www.tcd.ie/Nursing_Midwifery/

Appendix 8; Further Visual Inspection Correlational Studies

8.1 Visual Inspection Studies from Construction Engineering

Four correlational studies from the construction engineering domain used a methodology of participants being asked to observe targets, hazards or defects using graphical representations of workplaces. Even though these studies are laboratory based and simulate “real world” searches, they nevertheless present quantitative evidence as to how many hazards EHS professionals observed under lab conditions.

Perlman, et al, (2014) used 61 participants all with a construction education background. 23 participants were student engineers and 38 were construction managers. The participants were subjected to both virtual reality depictions and photographs of construction sites. The results demonstrated that participants identified between 50% and 99% of the 48 hazards observable within the images.

Dzeng et al. (2016) used four PC depicted construction site images with 25 participants. Ten of these participants were experienced construction workers and 15 were students studying for degrees in construction. They found that the level of visual observation varied from 80.39% to 86.71% of the hazards present in the images.

Dirksen, et al (2013) reviewed sewer defect reports, conducted by engineers using visual inspection. Using photographs contained within these reports the authors found that engineers missed defects in around 25% of visual inspections.

Carter & Smith, (2006) reviewed 90 method statements that contained identified hazards, from three large and complex construction sites. These authors listed hazards that were not identified in these method statements, by using expert scenario analysis. They found missed hazards in 33% of the method statements.

Kempton et al, (2000) filmed two houses to show front and rear building elevations to building surveyor participants. Each house was viewed by 38 surveyors and the outcome variable was estimated costs of repair. These scholars reported that surveyor variability in visual inspection was demonstrated. Furthermore they report ambiguity with this type of visual inspection stating that a house survey is not an exact science, but consists of subjective judgements as to condition. They further report that any variability in hazards observed is not necessarily due to error, but can be attributed to subjectivity.

8.2 Visual Inspection Studies from Industrial Engineering.

Gallwey, (1998b) presents date which is summarised into table 10 below and cites twelve studies from 1941-1982 involving VIP. The commonality of these studies centres on their application to industrial engineering, and specifically the visual inspection of the following industrial components. Again the evidence here is just how variable VIP is when visual inspections are required.

Table 28 Selected Industrial Sector VIP Findings

Item inspected	%age defects observed	Measure used
Sheets of tin plate	79	Detection of surface defects
Piston rings	10	Agreement of verdict between inspectors
Wiring and solder	80	inspection accuracy
Ball bearings	63	Detection of faults
Ceramic articles	36	Consistency of decisions
Coins	55	Fault detection rate
Printed circuit boards	73-95	Fault detection rate
Tin cans	58	Fault detection rate
Rubber seals	92	Fault detection rate
X-ray plates	20	Faults missed by radiologists
Knitwear products	50	Fault detection rate
Sauce pans and Dutch ovens	62-75	Fault detection rate

Rao et al, (2006) presents a study where twenty eight University students were trained in a real-world contact lens inspection task. This visual search was conducted in a laboratory using PC’s to display the contact lenses. The defects detected rate varied from 72%-80%.

Gramodpadhye et al, (2009) used 28 laboratory based university students trained in a real world contact lens inspection study. They found that the factors that affected VIP included defect type, probability, distribution and complexity of relevant standard. The found that a high number of defects in a search field negatively affected search performance. They also found that low defect probability combined with smaller number of defect types improved visual search but worsened decision making. Finally they found that when defect types and probabilities were increased, decision making improved but search performance decreased.

8.3 Visual Search Studies from Security Screening

Mitroff & Biggs, (2013) conducted an experiment using a smartphone app called Airport Scanner” a mobile phone game which simulated x-ray baggage inspection. They found that for ultra rare items, when fewer than 0.05% of cases containing the target were the subject of visual search by inspectors, the error rate was circa 45%.

Wolfe et al, (2005) conducted laboratory experiments where artificial baggage screen images closely resembling x-ray scanners. Observers looked for objects that were displayed within these baggage screen images. The number of objects “planted” within certain screens images where N= 3, 6, 12 or 18. The prevalence rate of screen images with these objects was set at 1%, 10% and 50%. The results showed that at 1% prevalence, the error rate of observers was 30%. At 50% prevalence, the error rate was 7%. These scholars summarise their findings by stating that ultra low prevalence rates for targets negatively affects visual inspection performance.

Wolfe & Horowitz, (2007) used university observers tasked with visually inspecting realistic depictions of x rayed baggage for weapons. They reported an error rate for missed weapons of between 0% and 40% with the higher miss rates corresponding to lower weapon prevalence.

8.4 Visual Search Studies from Medical Image Diagnostics.

Evans et al, (2013) reviewed the accuracy of radiologists observing cancerous lesions in medical scans. They found that if the prevalence rate in the scans were high (50% of images had cancerous lesions) then the miss rate was 12%. If the prevalence rate was low, (3%) the miss rate rose to 50%. In a related study by Drew et al, (2013) looking at the in-attentional blindness phenomenon, twenty four radiologists were tasked with searching medical scans. Only four of these radiologists noticed an image of a gorilla that had been inserted into the corner of all the scans.

8.5 Visual Inspection Failures that were Judicially Investigated

There have been seven Governmental enquiries published whereby a visual inspection failure had been identified as a contributing factor to fatalities or economic loss. These enquiries have academic weight, as they were conducted under rigorous investigative and judicial conditions with expert witness contributions. These cases are the Deepwater Horizon oil well blow out in the Gulf of Mexico, two US based nursing home fires, one UK based nursing home fire, three US air crash investigations and the interim report into the beef burger adulteration scandal in Ireland & the UK, whereby horse meat was found in food products.

Hopkins, (2010) reports that in the seven hours leading up to the Deepwater Horizon oil well blow out in the Gulf of Mexico in 2010, a group of four senior managers with extensive engineering experience visited the rig. Had they observed the mud monitoring data in the control room they would have seen that well pressure was a cause for concern. This overpressure would have been immediately raised with the rig manager and the impending disaster could have been averted. Instead the senior managers remained on the rig up until the explosion occurred and visually inspected other aspects of the rig operation, such as working at height provisions and slip trip & fall prevention.

The US Government Auditing Organisation investigated fire safety standards in American nursing homes in the aftermath of two fires that claimed the lives of over 40 residents (GAO, 2004). In each nursing home they found prior fire safety inspections had missed constructional defects that had contributed to the fatalities. In one specific example highlighted, the failure of the inspector to notice an opening in the ceiling connecting two floors had contributed to additional deaths from smoke spread.

Similarly in the Rosepark nursing home fire in Lanarkshire Scotland where 14 residents died, a prior inspection by a health and safety consultant was found to have missed many significant fire hazards (Lockhart 2011). The coroner's report concluded that; an adequate inspection would have identified the inappropriate storage of flammable aerosols in an electrical cabinet, which if rectified would have prevented the fatalities.

Drury, (2002) states that reliability issues with visual inspection caused the explosive decompression involved in the Aloha incident Hawaii 1988, and the complete engine failure in the Pensacola incident Florida 1997. See, (2012) refers to 111 fatalities from the Sioux City aircraft crash landing in 1989 and states that this crash was directly caused by a cracked engine fan disk, that was not detected during visual inspection.

A further but more generalised judicial finding into visual inspection failure involves the interim Elliot, (2013) report. This was commissioned by the UK Government in response to the horse meat adulteration scandal and found that food auditing quality was variable. Professor Elliot comments on one particular case are illustrative. One food auditor's conduct was so poor, that the company gave this auditor a list of minor non compliances out of pity.

8.6 Interpretivist Visual Inspection Studies

The following two studies by Woodcock, (2014) and Hrymak et al, (2015) present two studies of VIP amongst EHS professionals. Although qualitative in nature, they both point out the difficulties involved in visual inspection. The Woodcock, (2014) study involved qualitative data generated by interviewing and shadowing fairground ride inspectors and food safety inspectors over a two year period. Woodcock, (2014) found that inspectors varied in their approach to inspections, but they all reported difficulty with relevant standards due to interpretative ambiguity.

Hrymak et al, (2015) also showed the variability of visual inspection conduct with this study that involved 40 experienced Irish EHS professionals. When asked exactly how they conducted risk assessments and safety audits, the participants all reported using three distinct methods; observation, asking questions and reading on site documents. However when asked for the exact procedures involved, there was a great deal of variation. In terms of visual inspection conduct, there was little standardisation as to how to look, what to look for, in what order to look, where to go to look, or how long to stay looking. Although all practitioners reported using observation as part of their hazard identification methods, their self-reported visual inspection conduct varied greatly.

REFERENCES

Ainsworth, L. 1982 . An RSM investigation of defect rate and other variables which influence inspection. Proceedings of the Human Factors Society Annual Meeting, 26, 868-872.

Aimola-Davies, A. M., Waterman, S., White, C.R., Davies, M., 2013. When you fail to see what you were told to look for: Inattentional blindness and task instructions. Consciousness and Cognition. 22, 221-230.

Albert, A., Hallowell, M.R., Kleiner B.M., 2014. Enhancing Construction Hazard Recognition and Communication with Energy-Based Cognitive Mnemonics and Safety Meeting Maturity Model: Multiple Baseline Study. Journal of Construction Engineering Management. 140, 1-11.

Albert, A., Hallowell, M.R., Skaggs M., Kleiner B.M., 2017 Empirical measurement and improvement of hazard recognition skill. Safety Science, 93. 1–8

Alvarez, G., A., Cavanagh, P., 2004. The capacity of visual short-term memory is set both by visual information load and by number of objects. Psychological Science, 15(2), 106–111.

American Red Cross, 2007, Lifeguarding, Yardley, PA: Staywell

Aronson, E., 2012. The Social Animal. Eleventh Edition. Worth Publishers.

Asada H., Sotozaki T., Endoh S., Tomita T., 1998. Practical Evaluation of Crack Detection Capability for Visual Inspection in Japan. Chapter In; Research and Technology Organisation. Applied Vehicle Technology Panel, Airframe Inspection Reliability under Field/Depot Conditions.

Aven, T., 2011. Quantitative Risk Assessment, the Scientific Platform. Cambridge University Press.

Baddeley AD. 2007. Working Memory, Thought and Action. Oxford, UK: Oxford Univ. Press.

Bahn, S., 2013 Workplace hazard identification and management: The case of an underground mining operation. Safety Science 57 129–137.

Blewett, V., O'Keeffe, V., 2011. Weighing the pig never made it heavier: Auditing OHS, social auditing as verification of process in Australia. Safety Science, 49 1014–1021

Biggs A.T., Cain, M.S., Clark, K., Darling, E.F., Mitroff S.R., 2013(b). Assessing visual search performance differences between Transportation Security Administration Officers and non-professional visual searchers. Visual Cognition. 21: 330-352

Biggs A.T., Mitroff S.R., 2013. Different predictors of multiple-target search accuracy between non-professionals and professional visual searchers. The Quarterly Journal of Experimental Psychology.

Biggs A.T., Mitroff S.R., 2014a. Improving the Efficacy of Security Screening Tasks: A Review of Visual Search Challenges and Ways to Mitigate Their Adverse Effects. Applied Cognitive Psychology. Published online in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/acp.3083.

Biggs A.T., Mitroff S.R., 2014b. Differences in multiple target search performance between non-professional and professional searchers due to decision making criteria. British Journal of Psychology. (wileyonlinelibrary.com) DOI: 10.1111/bjop.12096.

Brener, J., Oostman, M. 2002., Lifeguards watch, but they don't always see! World Waterpark Magazine, 14–16. May.

BS EN 14001, 2004. Environmental management systems. Requirements with guidance for use. British Standards Institute.

Business in Ireland. 2012. Central Statistics Office. Stationery Office, Dublin, Ireland.

Carter, G., Smith S.D., 2006. Safety Hazard Identification on Construction Projects. Journal of Construction Engineering Management.132:197-205.

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):112-25.

Cain, M.S., Dunsmoor, J.E., LaBar, K.S., Mitroff S.R. 2011. Anticipatory Anxiety Hinders Detection of Second Target in Dual-Target Search.

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.

Cain, M. S., Vul, E., Clark, K., Mitroff, S. R. (2012). A Bayesian optimal foraging model of human visual search. Psychological Science. 23(9), 1047–1054.

Chan, A.H.S., Chiu, C.H.Y., 2010. Visual Lobe Shape Characteristics of Experienced Industrial Inspectors and Inexperienced Subjects. Human Factors in Ergonomics and Manufacturing and Service Industries. 20 (5) 367-377.

Charnov, E.L. 1976. Optimal foraging, the marginal value theorem. Theoretical Population Biology. 9, 129–136.

Clark, K.,Cain, M.S., Adcock, A.R., Mitroff S.R., 2014. Context matters: The structure of task goals affects accuracy in multiple target visual search. Applied Ergonomics. 45, 528-533

Clift, L., Lawton, C. Maguire, M., 2011. Checking the checklist: the effect of training on the application and effectiveness of checklist based risk assessments. Loughborough Design School. Report for the Institution of Occupational Safety and Health. Loughborough University Institutional Repository. Leicestershire, UK.

Cohen, J. 1992. A Power Primer. Psychological Bulletin, 1992, Vol. 112. No. 1,155-1,159.

Colman, A.M., 2009. The Oxford Dictionary of Psychology. 3rd Edition Oxford University Press.

Cresswell, J.W., 2013. Research Design. Sage Publications

Dalton, J., Drury C.G., 2004. Inspectors performance and understanding in sheet steel inspection. Occupational Ergonomics, 4,51-65.

Demichela, M, Camuncoli G., 2014. Risk based decision making. Discussion on two methodological milestones. Journal of Loss Prevention in the Process Industries 28, 101-108.

DiCarlo, J.J., Zoccolan, D., Rust, N. C., 2012. How Does the Brain Solve Visual Object Recognition? Neuron 73, February 9.

Dirksen, J., Clemens, F.H.L.R., Korving, H., Cherqui, F.,Le Gauffre, P., Ertl, T, Plihal, H., Müller, K., Snaterse, C.T.M. 2013. The consistency of visual sewer inspection data. Structure and Infrastructure Engineering. Vol 9, Issue 3, 214-228.

Douglas, E., Cromie, S. & Leva, M.C. 2015. Hazard Perception and Reporting. Ergonomics and Human Factors. Daventry, England, 13-16 April. doi:10.13140/RG.2.1.3526.0966

Drew, T., Vo, M.L.H., Wolfe, J.M., 2013. The Invisible Gorilla Strikes Again: Sustained Inattentional Blindness in Expert Observers. Psychological Science. 24 (9), 1848-1853

Drew. T., Wolfe, J.M., 2013. Hybrid search in the temporal domain: Evidence for rapid, serial logarithmic search through memory. Attention Perception Psychophysics. DOI 10.3758/s13414-013-0606-y Published online 17th December 2013.

Drew. T., Boettcher, S.E.P., Wolfe, J.M., 2015. Searching while loaded; Visual working memory does not interfere with hybrid search efficiency but hybrid search uses working memory capacity. Psychonomic Bulletin and Review 9th June 2015.

Drury, C.G., 2002. Visual Inspection Reliability: What We Need To Know and Why We Need To Know It. 16th Human Factors in Aviation Maintenance Symposium. April 2-4 Available from http://hfskyway.faa.gov Accessed April 2018.

Drury, C.G., Karwan, M.H., & Vanderwarker, D.R. (1986). The two-inspector problem. IIE Transactions, 18, 174-181.

Drury, C.G., Watson J., 2002. Good Practices in Visual Inspection. Federal Aviation Authority. US Department of Transportation. Washington DC. Available from http://www.faa.gov/about/initiatives/maintenance_hf/library/?q=visual+inspection. Accessed April 2018.

Drury, C.G., 1999. Human Reliability in Civil Aircraft Inspection. Paper presented at the RTO HFM Workshop on; The Human Factor in System Reliability. December 1-2. Available from http://hfskyway.faa.gov Accessed April 2018.

Dzeng, R.J., Chin-Ten, L., Yi-Cho, F., 2016. Using eye-tracker to compare search patterns between experienced and novice workers for site hazard identification. Safety Science 82 56-67.

Elliot, C., 2013. The Elliott Review into the Integrity and Assurance of Food Supply Networks- interim report. HM Government UK. Available at; https://www.gov.uk/government/publications.

Elliot, C., 2014. Elliott Review into the Integrity and Assurance of Food Supply Networks – Final Report A National Food Crime Prevention Framework HM Government UK. at; https://www.gov.uk/government/publications.

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

EU-OSHA, (2014) Psychosocial risks in Europe: Prevalence and strategies for prevention, Publications Office of the European Union, Luxembourg.

Evans KK, Birdwell RL, Wolfe JM (2013). If You Don’t Find It Often, You Often Don’t Find It: Why Some Cancers Are Missed in Breast Cancer Screening. PLoS ONE 8(5):

Field, A., 2013. Discovering Statistics Using IBM SPSS Statistics. 4th Edition. Sage Publications Ltd.

Fiske, S.T. & Taylor, S.E. 1991. Social cognition (2nd ed.). New York: McGraw-Hill.

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

Fleck, M.S., Samei, E., Mitroff, S.R., 2010. Generalized "satisfaction of search": adverse influences on dual-target search accuracy. Journal of Experimental Psychology: Applied 16 (1), 60-71.

Foot, H.C., & Russon, J.M. (1975). The reciprocation of unfavorable evaluations of performance in a two-person inspection task. European Journal of Social Psychology, 5, 289-296.

Forster, A.M., Douglas, J., 2010. Condition survey objectivity and philosophy driven masonry repair. Structural Survey Vol. 28 No. 5

Fox, J.G. 1973. Recent human factors contribution to enhancing industrial quality control. Behaviorometric, 3, 99-118.

Fougnie, D., Cormiea, S.M., Zhang, J., Alvarez, J.A., Wolfe, J.M., 2015. Winter is coming: How humans forage in a temporally structured environment. Journal of Vision, 15(11):1, 1–11

Gaglio, B., Phillips, S.M., Heurtin-Roberts, S., Sanchez, M.A., Glasgow, R.E., 2014. How pragmatic is it? Lessons learned using PRECIS and RE-AIM for determining pragmatic characteristics of research. Implementation Science. 9:96.

GAO, 2004. Nursing Home Fire Safety: Recent Fires Highlight Weaknesses in Federal Standards and Oversight. Report to Congressional Requesters, United States Government Accountability Office GAO-04-660. http://www.gao.gov/products/GAO-04-660 (Dec 14, 2015).

Gallwey T.J.,1998a Evaluation and Control of Industrial Inspection. Part 1. Guidelines for the Practitioner. International Journal of Industrial Ergonomics. 22, 37-49.

Gallwey T.J.,1998b Evaluation and Control of Industrial Inspection. Part 2. The scientific basis for the guide. International Journal of Industrial Ergonomics. 22, 51-65.

Gallwey, T.J., Drury, C.G., 1996. Task Complexity in Visual Inspection. Human Factors. (28), 5, 595-606.

Gerbec, M., Baldissone, G, Demichel, M., 2017. Design of procedures for rare, new or complex processes: Part 2 – Comparative risk assessment and CEA of the case study. Safety Science 100, 203–215.

Gawande, A. 2009. The Checklist Manifesto, How to Get Things Right. Profile Books

Geyer, L.H., Patel, S., Perry, R.F. 1979. Detectability of multiple flaws. Human Factors, 21, 7-12.

Ghylin, K.M., Drury, C.G., Schwaninger, A. 2006. Two-component model of security inspection: Application and findings. Paper presented at the 16th World Congress of Ergonomics, IEA 2006, Maastricht, The Netherlands.

Gilovich, T., Griffin, D.W.H., Kahneman, D., 2013. Heuristics and Biases, The Psychology of Intuitive Judgement. Cambridge University Press.

Gould, J., Glossop, M., Ioannides, A. 2005. A Review of Hazard Identification Techniques HSL/2005/58. Health and Safety Laboratory, HMSO. UK.

Gramopadhye, A.K., Desai, R.R., Bowling, S., & Khasawneh, M. (2003). Task analysis of general aviation inspection activities: Methodology and Findings. Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting, 47,36-40.

Gramopadhye, A.K., Drury C.G., Jiang, X., Sreenivasan, R., 2002. Visual search and lobe size: can training on one affect the other? International Journal of Industrial Ergonomics. 30, 181-195

Gramopadhye, A.K., Drury C.G., Sharit, J., 1997. Feedback strategies for visual search in airframe structural inspection. International Journal of Industrial Ergonomics. 19, 333-344.

Gramopadhye, A.K., & Madhani, K. (2001). Visual search and visual lobe size. In C. Arcelli et al. (Eds.), IWVF4, LNCS 2059 (pp. 525-531). Berlin: Springer-Verlag.

Gramopadhye, A.K., Melloy, B.J., Gopinath, M., Budgavi M,. 1997b. an evaluation of economic and performance feedback in an inpection task with explicit economic consequences. International Journal of Industrial Ergonomics. 20, 327-337.

Gramopadhye, A.K., Melloy, B.J., Him, H., Koenig, S., Nickles, G., Bingham, J., Kaufmann, J., Thaker, J., Master, R, 2000. Proceedings of the IEA 2000HFES 2000 Congress.

Hancock, P.A., 1984. Environmental stressors. In J.S. Warm (Ed.), Sustained Attention in Human Factors (pp. 103-142). Chichester: Wiley.

Harris, D.H, 1968. Effect of defect rate on inspection accuracy. Journal of Applied Psychology. 52,377-379

Heidl, W., Thumfart, S., Eitzinger, C., Lughofer, E., Klement, E.P., 2010. Classifier-based analysis of visual inspection: Gender differences in decision-making. Paper presented at the 2010 IEEE International Conference on Systems, Man, and Cybernetics, Istanbul, Turkey.

Health and Safety Review. March 2018. IRN publishers, Dublin 6.

Hillman, B.J., Swensson, R.G., Hessel, S.J., Gerson, D.E., & Herman, P.G. (1976). The value of consultation among radiologists. American Journal of Roentgenology, 127, 807-809.

Howick J., and Mebius, A., 2014. In search of justification for the unpredictability paradox. Trials, 15:480.

Hollis M., 2000. Surveying Buildings. RICS Books Coventry, England.

Hollis, M., Bright, K., 1999. Surveying the Surveyors. Structural Survey. Vol 17, No.2. 65-75

Hong, S.K. 2005. Human stopping strategies in multiple target searches. International Journal of Industrial Ergonomics. 35, 1-12.

Hopkins, A., 2011. Management walk arounds: Lessons from the Gulf of Mexico oil well blowout. Safety Science 49, 1421-1425.

Hrymak V., DeVries J.M.A., Leva M.C., 2015. Workplace Risk Assessment in Ireland; is there a standard practice. Proceedings of the Irish Ergonomics Society Annual Conference. Trinity College Dublin. June 2015.

HSA 2006. Guidelines on Risk Assessments & Safety Statements. Health and Safety Authority. James Joyce St, Dublin 1. Ireland.

HSA, 2015. Safety Representatives Resource Book Health and Safety Authority, James Joyce St, Dublin 1. Ireland..

HSA, 2017. Summary of Workplace Injury, Illness and Fatality Statistics for 2014-2015. James Joyce St, Dublin 1. Ireland.

HSE, 2014. Identify the Hazards. Available from; http://www.hse.gov.uk/risk/identify-the-hazards.htm. Accessed April 2018.

HSE, 2017. How Do I Do a Risk Assessment. Available from; http://www.hse.gov.uk/risk/faq.htm. Accessed April 2018.

Hutchison, D., Styles, B., 2010. A Guide to Running Randomised Controlled Trials for Educational Researchers. National Foundation for Educational Research, Slough Berkshire. UK.

International Labour Organisation. 2017. OSH Thesaurus. Available from; http://www.ilo.org/dyn/oshthes2/en/f?p=14660:1:::NO::: Accessed April 2018.

IEC/ISO 31010, 2009. Risk Management - Risk Assessment Techniques. International Standards Organisation. Geneva 20, Switzerland.

ISO 19011:20110 Guidelines for Auditing Management Systems. International Standards Organisation. Geneva 20, Switzerland.

ISO 22000, 2005. Food Safety Management Systems. International Standards Organisation. Geneva 20, Switzerland.

ISO 45001, 2018. Occupational Health and Safety Management Systems. International Standards Organisation. Geneva 20, Switzerland.

Jiang, X., Gramopadhye, A.K., Melloy, B.J., & Grimes, L. (2007). Correlating trust and performance in a hybrid inspection environment.

Jiang, X.J., Master B.R., Gramopadhye, A.K., Melloy, B.J., 2000. A system to understand human machine function allocation issues in visual inspection. Proceedings of the IEA 2000/HFES Congress.

Johansen, I.L., Rousand, M., 2015. Ambiguity in risk assessment. Safety Science 80, 243–251.

Juran, J., 1935. Inspector's errors in quality control. Mechanical Engineering 57 (1), 643–645.

Kane, J., Moore, D., Ghanbartehrani, S., 2009 The effects of expectations on visual inspection performance. Oregon State University, Research Paper

Kahneman, D., Klein, G., 2009. Conditions for intuitive expertise: a failure to disagree. American Psychologist. September, 515-526.

Kahneman, D., 2011. Thinking Fast and Slow. Farrar, Strauss and Giroux Publishers.

Kempton, J., Nicol, S, Anumba, C J and Dickens, J (2000) Analysis of surveyor variability from data obtained from surveys of dwellings represented on video film. In: Akintoye, A (Ed.), 16th Annual ARCOM Conference, 6-8 September 2000, Glasgow Caledonian University.

Kempton, J. Alani, A., Chapman, K., 2002. Potential effects of the confirmation bias in house condition surveys. Structural Survey Vol 20. No.1 190-197

Kempton, J. Chapman, K., Alani, A., 2002. Surveyor element lifetime predictipons and the anchoring and adjustment heuristic. Surveyor element lifetime predictions and the anchoring and adjustment heuristic. In: Greenwood, D (Ed.), 18th Annual ARCOM Conference, 2-4 September 2002, University of Northumbria. Association of Researchers in Construction Management, Vol. 1, 145-54.

Kempton, J., 2004. Stock condition data for decent homes: impact of surveyors' judgments. Structural Survey. Vol 22 No. 3. 126–130.

Lanagan-Leitzel, L.K., Skow, E., Moore, C.M., 2015. Great Expectations: Perceptual Challenges of Visual Surveillance in Lifeguarding. Applied Cognitive Psychology 29: 425–435.

Le Coze J.C., 2005. Are Organisations Too Complex For Technical Risk Assessment and Current Safety Auditing?. Safety Science 43 613-638.

Lenhardt U., Beck D., 2016. Prevalence and quality of workplace risk assessments – Findings from a representative company survey in Germany. Safety Science 86, 48–56.

Lockhart, B.A., 2011. Inquiry into the deaths of 13 Residents of Rosepark Nursing Home Uddingston Scotland. Sherrifdom of South Strathclyde Dumfries and Galloway. Scotland.

Lofstedt R. and Frewer L., 1998 " Risk and Modern Society" Earthscan Publications

Makin, M.A., Winder, C., 2008. A new conceptual framework to improve the application of occupational health and safety management systems. Safety Science 46, 935–948.

Marhavilas, P.K., Koulouriotis, D, Gemeni, V., 2011. Risk analysis and assessment methodologies in the work sites: On a review, classification and comparative study of the scientific literature of the period; 2000-2009. Journal of Loss Prevention in the Process Industries 24, 477-523.

Mariken, H.C., Everdij, H, A.P., Blom, A.P., 2013. Safety Methods Database Version 1.0 Maintained by NLR. Netherlands Aerospace Centre Available from http://www.nlr-atsi.nl/services/safety-methods-database/ Accessed April 2018.

Marshal, P., Hirmas, A., Singer M., 2018. Heinrich's pyramid and occupational safety: A statistical validation methodology. Safety Science. 101, 180-189

McCallum, M., Bittner, A., Rubinstein, J., Brown, J., Richman, J., Taylor, R. 2005. Factors contributing to airport screener expertise. Proceeding of the Human Factors and Ergonomics Society Annual Meeting, 49, 922-926.

Melloy, B.J., Harris, J.M., Gramopadhye, A.K., 2000. Predicting the accuracy of visual search performance in the structural inspection of aircraft. International Journal of Industrial Ergonomics. 26, 277-283.

Miller, G.A., 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information". Psychological Review. 63 (2): 81–97

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, Vol. 25(1) 284–289.

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. Vol. 2(1) 121–128.

Moher D, Liberati A, Tetzlaff J, Altman D.G., 2009. The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7):

Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., Elbourne, D., Egger, M, Altman D.G., 2012. RESEARCH METHODS & REPORTING, CONSORT 2010, Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials BMJ 2010;340:c869 doi: 10.1136/bmj.c869

Montibeller G., von Winterfeldt D., 2015. Cognitive and Motivational Biases in Decision and Risk Analysis. Risk Analysis, Vol. 35, No. 7.

Moore, M., Brent, P.E., Phares, D. Graybeal D., Rolander D., Washer, G. 2001. Reliability of Visual Inspection for Highway Bridges. US Department of Transportation. Federal Highway Department. McLean Virginia. United States of America.

Nakayama, K., Martini, P., 2011. Situating visual search. Vision Research, 51, 1526-1537.

Nalangula, D., Greenstein, J.S., Gramopadhye, A.K., 2006. Evaluation of the effect of feedforward training displays of search strategy on visual search performance. International Journal of Industrial Ergonomics. 36, 289-300.

Neathey, F., Sinclair,A., Rick,J., Ballard,J., Hunt,W., Denvir, A., 2006. An evaluation of the five steps to risk assessment. 2006. Research Report 476. Health and Safety Executive UK.

New Scientist, 2015. Why Things Fall Apart. Number 3034, 28-33. 15th August

Nickles, G.M., Melloy, B.J., Gramopadhye, A.K., 2003. A comparison of three levels of training designed to promote systematic search behaviour in visual inspection. International Journal of Industrial Ergonomics. 32, 331-339.

Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. Science 28 August 2015. Vol 349 Issue 6251

OSHA, 2016. Occupational Safety and Health Administration. OSHA Fact Sheets DEP FS-3783 08/2016. US Department of Labor Available at; https://www.osha.gov/OshDoc/data_General_Facts/factsheet-inspections.pdf Accessed April 2018.

Palmer, J., Verghese, P., Pavel, M., 2000. The Psychophysics of visual search. Vision Research 40, 1227-1268.

Peissig, J.J., Tarr, M.J., 2007. Visual Object Recognition: Do We Know More Now Than We Did 20 Years Ago?. Annual Reviews in Psychology 58:75–96

Petty, N.J., Thomson, O.P., Stew, G., 2012a. Ready for a paradigm shift? Part 1: Introducing the philosophy of qualitative research. Manual Therapy. 17, 212-274.

Petty, N.J., Thomson, O.P., Stew, G., 2012b. Ready for a paradigm shift? Part 2: Introducing qualitative research methodologies and methods. Manual Therapy. 17, 378-384.

Perlman, A., Sacks, R., Barak, R., 2014. Hazard recognition and risk perception in construction. Safety Science 64. 22–31.

Pinto, A., Ribeiro, R.A., Nunes, I.L. 2013. Ensuring the Quality of Occupational Safety

Rajsic, J., Wilson, D.E., Pratt, J., 2017. The price of information: Increased inspection costs reduce the confirmation bias in visual search. The Quarterly Journal of Experimental Psychology.

Ramachandaran, V.S., 2012.The Tell-Tale Brain, Unlocking the Mystery of Human Nature. Chapter 2; Seeing and Knowing. Windmill Books, London

Rao, P., Bowling, S.R., Khasawneh, M.T., Gramopadhye, A.K., Melloy, B.J., 2006. Impact of Training Standard Complexity on Inspection Performance. Human Factors and Ergonomics in Manufacturing. 16, (2) 109-132

Reason, J. 1997. Managing the Risks of Organisational Accidents. Ashgate Publishing Ltd. Aldershot, UK.

Reddin, P., 2016. Inspecting and Diagnosing Disrepair. Routledge Abingdon, Oxon

RICS, 2010. Practice Standards. UK Building surveys and technical due diligence of commercial property. 4th Edition Guidance Note GN23/2010. Royal Institute of Chartered Surveyors. UK.

RICS, 2002. Stock Condition Surveys. RICS Guidance Note. Royal Institute of Chartered Surveyors. UK.

Robson, L.S., Macdonald, S., Van Eerd. D., Gray, G.C., Bigelow, P.L, 2000. Something Might be Missing From Occupational Health and Safety Audits: Findings from a Content Validity of Five Audit Instruments. Journal of Occupational and Environmental Medicine. Vol 52, No.5, 536-543

RTO., 1998. Executive Summary in; Research and Technology Organisation. Applied Vehicle Technology Panel, Airframe Inspection Reliability under Field/Depot Conditions. Papers presented at the RTO Workshop. Brussels, Belgium, 13-14 May. North Atlantic Treaty Organisation. Neuilly Sur Seine, Cedex, France.

Schultz, K.F., Grimes, D.A., 2002. Generation of Allocation Sequences in Randomised Trials; Chance not Choice. THE LANCET Vol 359, Feb 9.

Schulz, K.F., Altman, D.G., Moher, D., 2010. Correspondence CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Medicine , 8:18

Scopus Content Coverage Guide Updated August 2017. www.elsevier.com/__data/assets/pdf_file/0007/69451/0597-Scopus-Content-Coverage-Guide-US-LETTER-v4-HI-singles-no-ticks.pdf Accessed April 2018.

See J. E., 2012. Visual Inspection: A Review of the Literature. Prepared by Sandia National Laboratories Albuquerque, New Mexico 87185, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

See, J.E., 2015. Visual Inspection Reliability for Precision Manufactured Parts. Human Factors. VOl 57 No. 8. 1427-1442

Simons, D.J., Chabris, C.F. 1999. Gorillas in our midst: Sustained inattentional blindness for dynamic events. Perception, 28, 1059-1074.

Solman, G.J.F., Hickey K., Smilek, D., 2014. Comparing target detection errors in visual search and manually-assisted search. Attention Perception Psychophysics. 76: 945-958

Spencer, F.W.,1996. Visual Inspection Research Project Report on Benchmark Inspections. Report No. DOT/FAA/AR-96/65. FAA Aging Aircraft NDI Validation Center. Sandia National Laboratories, Albuquerque, NM, U.S.A.

Sadasivan S., Gramopadhye, A.K., 2009. Technology to support inspection training in the general aviation industry: Specification and design International Journal of Industrial Ergonomics 39, 608–620.

Schwaninger, A., 2005. Increasing efficiency in airport security screening. WIT Transactions on the Built Environment, Vol 82, 405-416.

Tabachnick, B.G., Fidell, L.S., 2014. Using Multivariate Statistics. Pearson Educational Limited.

Taylor, J.C. (1990). Organizational context for aircraft maintenance and inspection. Proceedings of the Human Factors Society 34th Annual Meeting, 34, 1176-1180.

Thackray, R.I. (1993). Correlates of individual differences in non-destructive inspection performance: A follow-up study. Available from http://hfskyway.faa.gov Accessed April 2018.

Tomaschek, A., Lanfer, S.S.L., Melzer, M., Debitz, U., Buruck, G., 2018. Measuring work related psychosocial and physical risk factors using workplace observations: a validation study of the health workplace screening. Safety Science. 101, 197-208

Tomkins, N., 2010. First Principles in Employer Liability. Journal of Personal Injury Law (3), 131-138

US National Commission 2011. Deep Water The Gulf Oil Disaster and the Future of Offshore Drilling Report to the President National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling January 2011

Van Stolk C., Staetsky, L., Hassan, E., Kim, C.W., 2012. Management of occupational health and safety; and analysis of the findings of the European Survey of Enterprises on New and Emerging Risks (ESENER) Luxembourg: Publications Office of the European Union.

Vora, J., Santosh N., Gramopadhye, A.K., Duchowski, A.T., Melloy, B.J., Kanki, B., 2002. Using virtual reality technology for aircraft visual inspection training: presence and comparison studies Applied Ergonomics 33, 559–570.
Wade N.J., Swanston M.T. 2013. Visual Perception; an Introduction. Psychology Press, New York.

Wadsworth, E., Walters, D., 2014. The determinants of workplace health and safety practice in the UK. policy and practice in health and safety. Volume 12, Issue 2, 3-22.

Wales, A.W.J., Anderson, C, Jones, K.L., Schwaninger, A., 2009. Evaluating the two-component inspection model in a simplified luggage search task. Behavior Research Methods, 41 (3), 937-943.

Walker, H.K., Hall W.D. Hurst J.W., 1990 Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition. Butterworths; 1990.

Wang, M.J., & Drury, C.G. (1989). A method of evaluating inspector's performance differences and job requirements. Applied Ergonomics, 20, 181-190.

Wang, J.J.M., Lin, S.C., Drury, C.G., 1997. Training for Strategy in Visual Search. Inter national Journal of Industrial Ergonomics. 20,101-108.

Watanapa, A., Kaewkuekool, S., Suksakulchai, S. 2012. Influence of training with and without reward on visual inspector's performance in a 3 dimension model. Applied Mechanics and Materials, 110-116, 2911-2917.

White J.P.M. 1993. The Tort of Negligence and Employer Liability. 446-462 Oak Tree Press.

Wiener, E.L., 1975. Individual and group differences in inspection. In C.G. Drury and J.G. Fox (Eds.), Human Reliability in Quality Control (pp.101-122). London: Taylor & Francis.

Wolfe, J. M., 1996. Visual search. In H. Pashler (Ed.), Attention. University College London Press. London, UK.

Wolfe, J.M., Butcher, S., J., Lee, C., Hyle, M., 2003. Changing Your Mind: On the Contributions of Top Down and Bottom Up Guidance in Visual Search for Feature Singletons. Journal of Experimental Psychology: Human Perception and Performance Vol 29, No. 2, 483-502.

Wolfe, J.M., Cain, M. S., & Alaoui-Soce, A. (2017). Hybrid value foraging: How the value of targets shapes human foraging behaviour. Attention, Perception & Psychophysics.

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

Wolfe, J.M., Horowitz, T.S., Kenner, N.M., 2005. Rare targets are often missed in visual search. Nature 435, 439-440.

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, 1-17.

Wolfe J.M., Evans, K.K., Drew, T., Aizenman, A., Josephs, E., 2015. How Do Radiologists use the Human Search Engine? Radiation Protection Dosimetry (2015), pp. 1–8 doi:10.1093/rpd/ncv501.

Wolfe, J.M., Aizenman, A.M., Boettcher, S.E.P., Cain, M.S., 2016. Hybrid foraging search; Searching for multiple instances of multiple types of target. Vision Research. 119, 50-59.

Wolfe, J.M., Horowitz, T.S. 2017. Five factors that guide attention in visual search. Nature: Human Behavior.

Woodcock, K., 2003. Task challenges in the performance of amusement ride inspection. Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting, 47, 980-984.

Woodcock, K., Model of Safety Inspection. 2014. Safety Science 62, 145-156

Wu, C.C., Wolfe, J.M. 2016 Multiple event monitoring Cognitive Research: Principles and Implications.

Yeow, P.H.P., Sen, R.N. (2004). Ergonomics improvements of the visual inspection process in a printed circuit assembly factory. International Journal of Occupational Safety and Ergonomics, 10, 369-385.

Zhang, J., Gong, X., Fougnie, D., Wolfe, J.M., 2015. Using the past to anticipate the future in human foraging behaviour. Vision Research, 111, 66-74

Zwarenstein, M., Treweek, S., Gagnier, J.J., Altman, D.G., Tunis, S., Haynes, B., Oxman, A.D., Moher, D., 2008. Research methods & reporting. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008;337: a2390.