OPTIMAL CHARACTERISTICS OF INSERTED GRAPHIC OBJECTS IN STIMULATING CCTV OPERATOR VIGILANCE AND PERFORMANCE

FIONA DONALD

CHAPTER 1: INTRODUCTION

Human vigilance is critical to the successful operation of many technical systems that are important to the functioning of society. These include air traffic control, radar and sonar, X-ray applications, long distance driving, and closed circuit television (CCTV) monitoring (Donald, 2001). Operators working in these types of environments are required to sustain their attention for long periods of time and maintain high levels of alertness (Warm, Parasuraman, & Matthews, 2008). The attention requirements of this type of work impose considerable demands on information processing, attention resources and working memory (Parasuraman, 1979; Warm, Dember, & Hancock, 1996). Further, stress and fatigue commonly occur in vigilance intensive positions (Sawin & Scerbo, 1995; Szalma, Warm, Matthews, Dember, Weiler, Meier, & Eggemeier, 2004). Numerous factors impact on the detection rates of operators of systems requiring vigilance, and their performance is frequently less than optimal (Donald & Andrew, 2003; Donald, Andrew, & Landgrebe, 2007; Edkins & Pollock, 1997; Keval & Sasse, 2006; Wells, Allard, & Wilson, 2006). Aligning human performance and the technological potential of these systems represents a major design and implementation challenge (Keval, 2006).

CCTV surveillance is one of the contexts where vigilance is important. The social relevance of CCTV and the fact that it provides a rich context for studying vigilance and visual search makes it an ideal area for vigilance research. CCTV forms a base from which findings can be applied to other jobs involving visual analysis and vigilance. In the last twenty years CCTV has been used extensively as a tool for improving public safety and protecting critical infrastructure in many countries, such as the United Kingdom (UK), Europe, United States of America (USA), Australia and South Africa. It is used throughout society including town and shopping centres, industrial and commercial settings, transport systems, and prisons (Neil, Thomas, & Baker, 2007). In the future it will be used increasingly in process control (Donald, C., 2008).

1.1. The effectiveness of CCTV

It is difficult to estimate the amount of money invested in CCTV. However, government expenditure provides an indication of expenditure by one segment of society (Donald, C., 2008). Between 1999 and 2003, £170 million was spent on 684 CCTV schemes by the UK's Crime Reduction Programme CCTV Initiative (Gerrard, Parkins, Cunningham, Jones, Hill, & Douglas, 2007). This amount only includes town or city centre surveillance and excludes expenditure by councils, schemes that were not included in Government Grants, and initiatives in public and private sector industry and commerce. Therefore total expenditure for this period is far in excess of the figure given. In London, it is estimated that individuals may be recorded by at least 300 cameras on average on any day (Gerrard et al., 2007). This illustrates the widespread nature of CCTV in the UK and the high costs involved.

The use of CCTV has increased substantially in the USA since the threat of terrorism, and has been accompanied by subsidisation from the USA government (Cameron, Kolodinski, May, & Williams, 2008). In the first seven years after the attacks of 11 September 2001, $23 billion was allocated to state and local governments for enhancing security and law enforcement, including CCTV surveillance. In addition, a budget of $230 million for the installment and maintenance of CCTV systems has been awarded by the Department of Homeland Security (Frost & Sullivan, 2008). The American Public Transport Association envisaged spending $1 billion in the USA between 2008 and 2010 (Gorshkov, 2008). Increased expenditure on CCTV systems worldwide is envisaged with estimated revenue of $46 billion for 2013 alone (ABI Research Press Release, 2008).

The large amounts invested suggest confidence in the effectiveness of CCTV surveillance in deterring, monitoring and responding to significant events or incidents, and producing evidence. Widespread acceptance and implementation of CCTV indicates that 'there has been an implied assumption that it will work and deliver results because it is such a good technology' (Donald, C., 2008, p.1). This sentiment is echoed by authors such as Gill and Spriggs (2005) and Squires (2010). However, it is challenged by several reports questioning the effectiveness of CCTV surveillance. Researchers have obtained mixed findings regarding CCTV effectiveness (Greenberg & Roush, 2009; Keval & Sasse, 2006; Wells et al., 2006; Welsh & Farrington, 2002). A review of 44 reports evaluating CCTV effectiveness was conducted by Cameron et al. (2008). Most of the sites were in the UK. Findings indicated that 41percent of systems resulted in decreased crime levels after CCTV had been implemented, 43 percent found no significant change in crime levels, and 16 percent reported increased crime. Five of the evaluations were in the USA, none of which found a statistically significant reduction in crime. In Australia, it was found that CCTV in public areas assists in detecting violent crime but does not prevent crime (Wells et al., 2006).

Certain types of offences are more likely to be detected on CCTV than others, with property crimes having higher detection rates than personal crime (Carli, 2008; Phillips, 1999). Thus car theft is more likely than assault to be detected on CCTV. In some instances, the implementation of CCTV has resulted in the displacement of crime to other areas, or conversely, the diffusion of crime reduction benefits to surrounding areas (Carli, 2008). This creates difficulties in evaluating CCTV effectiveness. On balance, evaluations of CCTV effectiveness suggest that the value added by CCTV may not be commensurate with the finances invested or expectations regarding results (Squires, 2010) although CCTV does assist in reducing crime and other significant events on a daily basis (Donald, C., 2008).

Numerous factors influence the effectiveness of CCTV systems and a thorough knowledge of these is needed in order to appreciate the complexities of CCTV effectiveness, diagnose issues, and design interventions. The current research needs to be understood in the context of this broader framework. Factors influencing CCTV effectiveness are categorised into five main components – the socio-political and industry contexts, organisational context, system and job design, and operator processes (Donald, F., 2010).

Amongst the reasons for the disappointing results of CCTV are the lack of coordinated planning and implementation strategy (Green, 2008), piecemeal implementation (Gerrard et al., 2007), inadequate partnerships with relevant stakeholders such as the police (Carli, 2008), technological shortfalls like poor image quality (Gerrard et al., 2007), proprietary systems that are incompatible with each other (Gerrard et al., 2007); management and organisational issues (Gill & Spriggs, 2005), control room design (Keval & Sasse, 2006), lack of objective crime analysis and intelligence (Corkill, 2006), and operator performance and search strategy (Keval & Sasse, 2008). In addition, CCTV has at times been implemented in situations where it is not appropriate or where its purpose is not clear (Gill & Spriggs, 2005).

As Squires states: ‘CCTV is but a tool, and where it was perceived to have failed this was often because the expectations placed upon it were too ambitious or because it was being used in unsuitable places for inappropriate problems’ (2010, p. 7).

The numerous factors that contribute to the effectiveness of CCTV systems indicate the complexity of the issues and the challenges in enhancing its effectiveness (Gill & Spriggs, 2005). Ultimately, a coordinated set of interventions is required to address the various aspects. However, the current research focuses on one aspect of CCTV effectiveness – the detection process conducted by CCTV surveillance operators. Operators are at the heart of the CCTV system and detection process, yet relatively little research has focused on them (Keval & Sasse, 2008). Neglect of the human side of CCTV may be based on the assumption that ‘the technology will work merely because it has the capability’ (Donald, C., 2008, p. 10).

The focus on operators is supported by Wells et al.’s (2006, p. 95) statement that ‘the effectiveness of CCTV may be very much dependent on a whole range of issues but in particular the monitoring strategies adopted by camera operators.’ Several other authors have also commented that the human side of CCTV has largely been ignored (e.g., Keval & Sasse 2006; Neil et al., 2007). In the current research, these personnel are referred to merely as ‘operators’ although they hold a variety of job titles in the operational context, such as security guard or officer, surveillance or control room operator, and review and analysis officer.

To detect significant events, operators need to be vigilant and deploy attention to displays. This is an implicit expectation of many CCTV system designers (Clarke, 1997). Yet the few studies that have focused on operators found that they frequently perform a variety of activities that remove their attention from the displays (Gill & Spriggs, 2005; Keval & Sasse, 2008; Smith, 2004; Tullio, Huang, Wheatley, Zhang, Guerrero, & Tamdoo, 2010). The assumption that operators continually and carefully monitor displays is therefore unlikely to hold true. When attention is deployed to displays, events that are easily visible are more likely to be detected than others (Wells et al., 2006). In addition, the majority of vigilance research conducted over the last 60 years indicates that vigilance performance decreases from 20 to 35 minutes (Sawin & Scerbo, 1995). Anecdotal evidence suggests that operator attention drifts after fifteen minutes (Security Park, 2001), but this has not been linked to detection performance. Together, these findings suggest shortfalls in the detection process and indicate the need for an intervention that assists operators in detecting significant events.

1.2. Focus of the current research

The main aim of the current research is to develop an intervention to enhance the detection levels of operators involved in proactive or real time CCTV surveillance. Proactive, live monitoring involves active search for significant events in real time, and is contrasted with reactive monitoring which involves reviewing previously recorded video in order to identify relevant events (Wells et al., 2006). The intervention consists of electronically inserting graphic objects or images into the video stream with the aim of assisting operators in detecting actual significant events.

Vigilance is crucial in the performance of operators and forms the overarching theoretical context for the research. However, extensive research into vigilance over many decades, the development of vigilance theories and the identification of numerous factors related to performance, have led to relatively few methods for enhancing vigilance being proposed by the mid-1980's (Mackie, 1987). Two of the methods that have been used are 'artificial signal injection' (ASI) and 'threat image projection' (TIP). The intervention developed in the current research is based on these interventions. ASI was developed as an aid to inspection tasks (Wilkinson, 1964). This involved placing a faulty item on the inspection line and assessing whether the inspector detected the fault. TIP is a similar and more recent intervention, developed and applied in the aviation industry for X-ray baggage screeners (Neiderman & Fobes, 2005). Here, an image of a potential threat, such as a weapon, is inserted electronically into the X-ray display of real baggage. The screener is required to detect the threat and has no way of knowing in advance whether the image is real or projected.

ASI and TIP form the basis for the intervention developed in this research. The intervention consists of electronically inserting graphics (inserted graphic objects or IGOs) into the CCTV stream which operators are required to detect and respond to. It shares some similarities with TIP in baggage screening but the differing context places different demands on operators resulting in an intervention with different qualities. However, it is noted that a key difference between ASI, TIP and the current intervention is that the detection of inserted objects is not an end in itself, but is intended to assist the detection of real significant events. This is a significant departure from ASI and TIP research, where the emphasis is more on the detection of the inserted objects themselves. IGOs, on the other hand, are intended to assist operators in engaging visually with the display, thereby assisting in detecting real significant real events.

The IGO intervention was chosen over the myriad of other possible interventions such as training, selection, automation, movement detection and face recognition software for several reasons. First, IGOs can be inserted on an on-going basis while operators perform their jobs. This is likely to reduce issues encountered with training courses regarding the generalisation of learning to the job context. Second, it will be argued that the IGOs are likely to assist in maintaining alertness and reducing potential vigilance decrements. Third, TIP has been used effectively with X-ray baggage screeners (Foulkes, Crombie, McClumpha, & Shadrake, 1997), suggesting that a similar intervention might have potential for the CCTV context. Last, IGOs and CCTV are seen as providing a useful platform for studying attention processes in the context of a complex, dynamic real world task. However, other interventions, such as those mentioned above, are also seen as being important and require future research.

The research focuses specifically on the design of IGOs. IGOs could potentially depict any objects, behaviours, and processes, and could represent these objects and events in a number of ways. For instance, they could be static or dynamic, visually conspicuous or difficult to perceive, represent important aspects of significant events or objects that are unrelated to the detection task. However, the types of objects and events they depict and how these are represented is likely to influence their ability to assist operators in detecting significant events. This is based on research findings that image design and characteristics influence whether images communicate the information they are intended to convey, and how easily users understand them (Cho, Ishida, Oyama, Inaba, & Takasaki, 2008). This suggests that IGO characteristics influence whether they are detected and understood by operators.

In the current research it is argued that IGOs that depict significant events are likely to create attention sets for the significant events depicted, and these in turn are likely to assist detection. This is based on previous vigilance and visual search research that found that knowledge of the target or significant event assists search (Davies & Parasuraman, 1982). This is important for CCTV, where significant events are frequently difficult to perceive and recognise, vary in nature, and are often deliberately camouflaged by perpetrators. Therefore it is important to develop operators' ability to perceive and recognise behaviours, objects, patterns and events that are associated with significant events. The current research therefore identifies a range of image characteristics in the literature and selects one characteristic for empirical study. The characteristic chosen is semantic distance, which refers to how closely IGOs depict target behaviours (TBs) that are associated with significant events. It will be argued that IGOs that depict TBs assist in the detection of significant events. On the other hand, IGOs whose content is unrelated to significant events are likely to have a negative effect on the detection of significant events as they could distract operators and are unlikely to assist in the creation and maintenance of attention sets that are primed for significant events.

The current research draws on literature from various fields to propose how IGOs could improve detection rates for significant events and ultimately contribute to CCTV effectiveness. These fields include vigilance, visual search, attentional capture and pictogram design. Much of the literature pertaining to the detection process is underpinned by attention processes. The application of theory to the CCTV detection process and intervention is a response to the challenge of translating effects that have been demonstrated in well-controlled laboratory experiments into effective designs and interventions in complex and dynamic real world settings (Wickens & McCarley, 2008), such as CCTV.

In addition to developing the IGO intervention, the current research also examines the level of detection performance on a specific vigilance task consisting of a ninety-minute video based on CCTV footage. Similarly, the existence of a vigilance decrement amongst operators for the video is examined. These establish the level of performance for this task and the need for an intervention to improve performance.

1.3. Structure of the thesis

The following chapter (chapter 2) focuses on vigilance performance issues and theories, and their implications for CCTV. This provides a broad understanding of vigilance research. The basis for the proposed intervention is explained more fully in chapter 3, as well as the intervention itself and how it differs from ASI and TIP. In chapter 4, the attention processes whereby IGOs could assist detection are explored. Research regarding attention guidance, scanning, situation awareness, mental models and attention sets is applied to the intervention. The chapter ends with a discussion of the potential negative implications of IGOs for attention resources, mental workload and distraction. These are discussed in relation to the need to reduce uncertainty in vigilance tasks. The specific focus of this research is on identifying IGO characteristics that could potentially assist the detection of real significant events. These characteristics are explained in chapter 5 where a range of possible IGO characteristics are discussed and certain of these (semantic distance and salience) are selected for the empirical side of the current research. The chapter ends with a statement of the research hypotheses. The methodology is described in chapter 6. In addition to the research design, sample, procedures and analysis methods used, this chapter describes the development of the materials used in the research. The results are presented in chapter 7. This includes the results for the operator sample as well an additional sample of novices. The results, their implications, and the theoretical contributions of the research are discussed in chapter 8. This chapter includes an overview of the limitations of the research and recommended future research directions.

1.4. Abbreviations used

For convenience, the following table lists the non-standard abbreviations used.

Table 1. Abbreviations

Abbreviation	Description
TB	Target behaviour
FA	False alarm
IGO	Inserted graphic object
RIGO	Inserted graphic object with random semantic distance
SIGO	Inserted graphic object with small semantic distance
Phase A	First thirty-minute phase
Phase B	Second thirty-minute phase
Phase C	Last thirty-minute phase
Nov	Novice
Gen	Generalist
Spec	Specialist
Surv back	Surveillance background (i.e., novice, generalist, specialist)
No. disengagement	Estimated number of disengagement periods per participant
Length disengagement	Estimated total length of disengagement period(s) per participant

CHAPTER 2: VIGILANCE THEORIES AND CCTV OPERATOR PERFORMANCE

Vigilance is a crucial element for system reliability and human safety in human-machine systems (Warm et al., 2008) and provides the overarching framework for the current research. This chapter defines vigilance, discusses selected theories that have been applied to vigilance and examines the trends emerging from vigilance research. Various definitions of vigilance have been proposed, with early writers understanding vigilance as a physiological state. Head (1926, p. 361) defined it as 'a state of high grade efficiency of the central nervous system.' This is consistent with neurologists and physiologists who use the term to refer to 'physiological efficiency' or 'arousal' (Parasuraman, 1984). Although arousal may play a role in vigilance, it is a general state that affects various attention processes in addition to vigilance. Early writers also emphasised the state of alertness required to respond to small changes which occur randomly in the environment (e.g., Mackworth 1957). These changes are referred to as signals, stimuli, targets, incidents, situations or significant events by different authors and these terms are used interchangeably.

More recent definitions focus less on these aspects and more on the maintenance of concentration or attention over long periods of time and the nature of the significant events that are observed. This is evident in the following definition: 'When attention must be focused on a source in order to detect a critical but infrequent event, this latter variety of attention is referred to as vigilance' (Gale & Christie, 1987, p. 164). Vigilance typically refers to situations where significant events must be detected within a context of neutral events (Craig, 1987; Lanzetta, Dember, Warm, & Berch, 1987; Sawin & Scerbo, 1995). In addition, the significant events are usually assumed to appear intermittently, infrequently and unpredictably. Depending on the work context, significant events may occur over hours, days, months or even years (Parasuraman, 1984).

Definitions have emphasised perceptual processes involving the simple detection of significant events. One such definition is that vigilance is 'sustained attention over a period of time (from seconds to hours) for targets that appear infrequently and unpredictably' (Parasuraman, 1984, p. 243). Similarly, Craig (1987, p. 645) stated that vigilance has traditionally been defined as 'sustained monitoring for an infrequent signal event.'

However, these definitions focus on the act of detection and lack a sense of readiness and urgency and the need for a response – aspects that are apparent in Mackworth's (1957) definition.

The above definitions overlook the complexity of the significant events and the cognitive processes required in many real world situations (Koelega, 1996). In particular, Donald (2001, p. 36) defines vigilance somewhat differently as 'a capacity for sustained effective attention when monitoring a situation or display for critical signals, conditions or events to which the observer must respond.' This definition omits the low frequency of significant events on the basis that certain vigilance intensive jobs (e.g., air traffic controllers) may involve frequent significant events but still require sustained monitoring, attention and performance. Further, the nature of the significant event is expanded to include greater complexity. It acknowledges that in addition to discrete and possibly simple events, significant events may include conditions that consist of patterns of events or deviations from norms that may not be immediately apparent.

Donald's (2001, p. 36) definition also states that 'Incorporated into this perspective on vigilance is the ability to identify, recognise and interpret the information that is being monitored'. This allows for cognitive processes that extend beyond the sensory recognition and detection of situations and is more relevant to a variety of jobs requiring vigilance. In addition to other cognitive processes involved in many vigilance intensive jobs, an understanding of the relevant context is often required in order to recognise and evaluate visual images accurately.

The terms 'vigilance' and 'sustained attention' are generally used interchangeably (e.g., Broadbent & Gregory, 1963; Jerison, 1977; Koelega, 1996; Lanzetta et al., 1987; Mackworth, 1970; Parasuraman, 1984). However, roles that require vigilance differ from those requiring only sustained attention (e.g., assessing student's assignments) in that for vigilance an immediate response is required to certain altered or unusual events. This is often accompanied by a sudden and large transition in mental workload (Huey & Wickens, 1993). Further, there is frequently a risk to life or property in failing to detect the event. Tasks that require vigilance differ from those that are merely repetitive or monotonous and can be distinguished by the following criteria:

‘1. The task should require detection, i.e. perceiving and reporting a change in operating environment.

2. The intensity of the signal should be close to the observer’s detection threshold, but the signal should be clearly perceivable when the observer is alerted or directed to it.
3. Signals should occur irregularly, infrequently, and if nonsignal stimuli are present, the ratio of nonsignals to signals should be high.
4. The task should be prolonged and continuous.’ (McGrath, 1963, cited in Koelega, 1996, p. 278-9).

Repetitive and vigilance tasks are not the same. Therefore the term ‘vigilance’ rather than ‘sustained attention’ is used in this study. However, attention processes from an integral part of vigilance and are incorporated into subsequent discussions. Donald’s (2001) definition of vigilance will be adopted for the purposes of this study because of its match with the CCTV surveillance context where SA is needed and complex cognitive and perceptual processes are involved in detection.

2.1. Performance Issues Related to Vigilance

There are two key issues regarding vigilance performance – the vigilance decrement and the overall level of vigilance. The former is a decrease in performance over time (Parasuraman, 1984) and is evidenced by slower detection times and lower detection rates over time resulting in decreased efficiency (Lanzetta et al., 1987). The decrease in performance generally occurs during the first 20 to 35 minutes, depending on the task being performed, and then stabilises at a significantly lower level (Sawin & Scerbo, 1995). In some laboratory experiments, reduced performance has been found after only five minutes of observation (Nuechterlein, Parasuraman, & Jiang, 1984). The vigilance decrement is attributed to the need to listen or look for a significant event for a prolonged time (Dember & Warm, 1979). The challenge is to develop interventions that reduce or prevent the vigilance decrement.

Vigilance tasks have long been recognised as challenging despite their frequently simple appearance. Numerous factors contribute to the vigilance decrement, including high mental workload and stress (Temple, Warm, Dember, Jones, LaGrange & Matthews, 2000; Warm et al., 2008), decreased conspicuity of significant events (Becker, Warm, Dember, & Hancock, 1991), high event rates, spatial and display uncertainty, noise (Parasuraman & Mouloua, 1996) and frustration (Warm, Dember, & Hancock, 1996). These conditions are present in the CCTV context, suggesting that the workload of operators is high. In addition, the complexity of CCTV displays and the viewing process (Donald, F., 2008) are likely to contribute to the operators' workload. Parasuraman and Mouloua (1996, p. 196) take this further by concluding that 'vigilance tasks not only deplete resources, they also drain the wellspring from which they come.'

The vigilance decrement has been demonstrated in numerous laboratory experiments (e.g., Craig, 1985; Mackworth, 1957; Parasuraman, 1979, 1986; Parasuraman, Warm, & Dember, 1987; Schmidtke & Micko, 1964; Warm, 1984; Wiener, 1987) and in studies using real life situations (Pigeau, Angus, O'Neill, & Mack, 1995) and simulations (e.g., Hollenbeck, Ilgen, Tuttle, & Sego, 1995; Molloy & Parasuraman, 1996). However, some studies have not found a vigilance decrement in complex tasks (e.g., Adams, Stenson, & Humes, 1961; Moray & Haudegond, 1998; Molloy & Parasuraman, 1996). These studies involved a mixture of laboratory and simulation-type tasks. Decrements in performance may be less severe in some operational contexts compared to laboratory contexts (Pigeau et al., 1995), in part due to measures that are already in place such as rest breaks and the interspersion of monitoring with other tasks such as patrolling or report writing.

In some operational contexts managers are already implementing strategies to prevent the vigilance decrement from occurring, for example, by rotating workers amongst tasks during shifts (Donald, 2001). The performance of other tasks simultaneously with monitoring, may also reduce the effects of monotony and provide more sensory input. Operators usually have the freedom to break monotony through various actions, such as eating, conversations with colleagues and moving around (Smith, 2004). These aspects are likely to reduce the vigilance decrement.

Regardless of the size or frequency of performance decrements, vigilance decrements may have significant implications for the safety and security of people and property, as Adams (1987, p. 738) points out:

> The issue is not the occasional occurrence of vigilance decrements in practical situations but whether they occur in important places. Are there for example, damaging decrements that degrade system-relevant performance in major job categories such as all radar operators or all fighter pilots, or in a major system such as the B-1 bomber?

This comment applies to CCTV where the failure to detect a significant event could result in damage to property and loss of lives. Detection failures do not, however, only result from the vigilance decrement. Operators' overall ability to detect significant events is equally important. Operators with a steady but low detection rate are also likely to miss significant events. This leads on to the second performance issue which is different from the vigilance decrement - the overall level of vigilance (Parasuraman, 1984).

The overall level of vigilance is a steady state that the observer brings to the task (Davies & Parasuraman, 1982). Thus a person with a high level of vigilance is likely to perform better on a detection task than one with a low level of vigilance. This has implications for the selection of operators. People with either high or low overall levels of vigilance may experience the vigilance decrement. However, the overall level of vigilance has not been clearly defined and it is not clear whether it is a personality or motivational construct, skill, cognitive capacity, or combinations of various cognitive processes. These cognitive processes could include, for example, the ability to self-regulate attention, large working memory capacity and effective supervisory executive control. Similarly, the interaction between overall levels of vigilance and the vigilance decrement are unknown. Despite the importance of both the vigilance decrement and the overall vigilance level, most research has focused on the former. The current research contributes to theory regarding the overall level of vigilance through qualitative analysis of interviews of participant responses and observation of participants during the vigilance task.

Researchers' strong emphasis on the vigilance decrement tends to assume that the initial level of detection is satisfactory, and that detection rates are only of concern once the decrement starts to occur. The current research questions this assumption by establishing the detection rate for a particular vigilance task and evaluating whether performance near the beginning of the task is satisfactory. In addition, the existence of a vigilance decrement for participants on the designated task is examined.

2.2. Vigilance Theories

Several theories of vigilance have been proposed, drawing strongly on motivation, habituation, inhibition, arousal, expectancy, signal detection and resource theories (Davies & Parasuraman, 1982; Huey & Wickens, 1993). Vigilance theories generally attempt to explain vigilance decrements rather than overall vigilance levels (Parasuraman, 1984), but are referred to as 'vigilance theories' in the literature.

Early researchers explained the vigilance decrement through learning and extinction (Mackworth, 1950), habituation (Broadbent, 1953), inhibition-reinforcement (McCormack, 1959) and operant conditioning (Holland, 1958). These concepts are similar to fatigue which can be overcome to an extent through rest pauses, alternating periods of observation with other tasks, and providing feedback on performance. However, mixed support was found for these explanations (Davies & Parasuraman, 1982). The effects of signal frequency, knowledge of results, rest pauses, inter-signal interval duration (Davies & Parasuraman, 1982) and the timing and circumstances under which the performance decrement sets in (Schmidtke, 1976) are not explained by these theories. Therefore the following discussion focuses on four selected theories that have implications for the intervention in the current research – arousal, signal detection (SDT), mindlessness, and multiple resource attention theories.

2.2.1. Arousal theory.

Arousal theories emphasise the general state of the individual as a predictor of performance on vigilance tasks (Davies & Parasuraman, 1982). Arousal has been described in both physiological and behavioural terms (Koelega, 1996). Many definitions have been proposed, focusing on the intensity of behaviour, motivational intensity and the individual's level of alertness, the last of which is most applicable to vigilance (Davies & Parasuraman, 1982).

Alertness includes the individual's readiness to react and ranges from sleepiness to high excitement (Berlyne, 1960, cited in Davies & Parasuraman, 1982).

Explanations of the vigilance decrement based on arousal theory have undergone various changes over time. Initially it was believed that vigilance tasks entailed low mental workload, with low levels of stimulation leading to under arousal and poor performance (Temple et al., 2000). Mental workload is defined as 'the interaction between the structure of systems and tasks on the one hand, and the capabilities, motivation, and state of the human operator on the other' (Kramer, Trejo, & Humphrey, 1996, p. 141). The assumption that vigilance tasks involve low mental workload was not based on empirical evidence and was challenged by the finding that these tasks' mental workload is high (Parasuraman & Mouloua, 1996) and that vigilance tasks are stressful (Warm et al., 2008).

More recent conceptualisations of arousal theory propose that an optimal level of arousal is associated with optimal performance (Loeb & Alluisi, 1984). Deviations from optimal arousal create a decline in performance, producing an inverted U-shaped curve. Performance is best at medium rather than high or low arousal levels (Koelega, 1996). It is hypothesised that monotonous and repetitive vigilance tasks progressively reduce arousal of the central nervous system (Schmidtke, 1976).

Empirical support for the relationship between arousal and vigilance performance is mixed (Koelega, 1996; Warm, 1977). The basic premise that arousal of the nervous system changes over time is generally accepted (Davies & Parasuraman, 1982). However, this is not restricted to tasks that require vigilance (Parasuraman, 1984). Where arousal decreases during vigilance tasks, there is not always a simultaneous change in detection efficiency, suggesting that arousal is not a key factor in the vigilance decrement (Koelega, 1996).

Arousal theory fails to take into account other aspects that affect vigilance performance, such as changes in expectancies and the response criterion (Loeb & Alluisi, 1984). Similarly, it does not explain vigilance decrements that occur under different conditions, such as those occurring when several tasks are undertaken concurrently (Loeb & Alluisi, 1984). Here, increased mental workload and distraction may remove any advantages gained by increased arousal. In addition, when arousal levels vary under different conditions, the decrement remains consistent instead of varying with arousal.

Changes in the conceptualisation of arousal over time provide for a more meaningful use of the concept. Initially seen as a unitary construct, arousal is now considered multi-dimensional (Matthews & Davies, 1998). Types of arousal include tonic versus phasic, and energetic versus tense arousal (Thayer, 1989). Tonic arousal refers to cortical arousal indicated by changes in brain-wave activity, as evident in states such as insobriety, intoxication or sleepiness (Parasuraman, 1984). Phasic arousal refers to physiological alterations such as changes in body temperature and heart rate, is shorter-term and influences how information is evaluated and used (Alluisi Coates, & Morgan, 1976). Energetic arousal ranges from vigour to fatigue and tense arousal from tension to relaxation (Thayer, 1989).

The distinction between energetic and tense arousal is particularly useful for vigilance, with energetic arousal making a significant contribution (Matthews & Davies, 1998). Energetic arousal is most similar to the concept of 'alertness' used in this research. In arousal theory terms, the intervention in the current research intends to maintain energetic arousal over time in order to assist the detection of significant events. This is done by inserting graphic objects into the video stream and requiring operators to respond to them. Based on arousal theory, this is likely to reduce monotony and provide more frequent stimuli to respond to.

In summary, the concept of arousal makes sense in so far as a continuum of alertness from sleepiness to excitement can be observed. For arousal to remain a useful concept, it is necessary to acknowledge changes in conceptualisation. At this point in the development of arousal theory, it is generally accepted that arousal is multi-dimensional and affects different resources in different ways. Consequently the relationship between arousal and performance is not necessarily represented by the inverted-U curve. In addition, people attempt to direct cognitive resources effectively, thereby compensating for their arousal levels rather than being passively involved in arousal processes (MacLean, Aichele, Bridwell, Mangun, Wojciulik, & Saron, 2009). By adopting a unitary view of arousal and assuming a direct link between central brain states and changes in performance, early arousal theories failed to consider the mediating role of information processing (Matthews & Davies, 1998), including expectancies and attention resources. Signal detection theory (SDT) and multiple resource attention theory provide insight into these aspects.

2.2.2. Signal detection theory.

Signal detection theory (SDT) has been applied to vigilance and provides many of the concepts and measures used in vigilance research (Long & Waag, 1981). In addition, it contributes to the current research by distinguishing between physiological and decision-making processes and highlighting the different roles these aspects play in detection. This is useful because ‘seeing’ or perceiving a significant event does not necessarily mean that the observer will decide that a significant event is present. This is illustrated in a study of X-ray baggage screeners where sixty percent of detection errors were errors in decision making rather than resulting from an inability to perceive the threats (Security General, 2002). The screeners could ‘see’ the threats and false alarms (FAs) but made erroneous decisions about them (Security General, 2002).

Similar issues are likely to exist in the CCTV context and may be aggravated by the low conspicuity and frequency of signals or significant events. Similarly to vigilance tasks, additional costs are likely to be imposed by the mere fact of looking continuously at CCTV displays as operators are unlikely to sustain attention for extended periods (Moray, 1984, Smith, 2004). SDT is a statistical decision theory and refers to the significant event to be detected as a signal or target (Davies & Parasuraman, 1982).

The measures used in vigilance research are typically based on SDT. These include hits (a correct response to a stimulus as a signal), FAs (an incorrect decision that a stimulus is a signal), correct rejections (decisions that non-signal stimuli are not significant events), and misses (incorrect responses to signals as non-signals) (Wickens & Hollands, 2000). These measures are applied in the current research. Good detection performance consists of maximising the number of hits and correct rejections and reducing the number of misses and FAs (Wickens & Hollands, 2000).

SDT is applicable to situations where a weak signal needs to be detected against a background of stimuli, or noise (MacMillan & Creelman, 1991). It proposes two stages during detection. Sensory information is gathered regarding the presence or absence of a stimulus. Thereafter, a decision is made regarding whether the sensory evidence indicates that a signal is present or not. These processes give rise to the concepts of sensitivity and the response bias (Balakrishnan, 1998).

Where there is a large change in the environment, detection is simple (Wickens & Hollands, 2000). However, when the event is close to the threshold of perception, detection is more difficult. Noise also increases detection difficulty and arises from external (such as the environment) or internal sources (e.g., neural 'noise' in sensory channels such as mental distractions), even when no signal is present. In the CCTV context, some significant events are detected more easily than others (Donald, 2001). Noise arises from multiple cameras competing for attention, image quality, stimuli in the image not related to significant events, and other sources. Operators may miss signals because of difficulties in separating signals from noise.

Noise affects the sensitivity of the detection mechanisms. Sensitivity refers to 'the keenness or resolution of the detection mechanisms' (Wickens & Hollands, 2000, p. 25), or sensory and perceptual processes involved in detection. It is the 'observer's ability to discriminate' between signals and non-signals and is related to stimulus characteristics' (MacMillan & Creelman, 1991, p. 8). Sensitivity represents the encoding of data (Balakrishnan, 1998). External stimuli create neural activity in the brain, with more intense stimuli generating more neural activity. The activity should exceed a perceptual threshold for the observer to decide that a signal is present (Wickens & Hollands, 2000).

A decrease in sensitivity over time has been found during vigilance tasks (See, Howe, Warm, & Dember, 1995). The most severe sensitivity decrements occur in successive tasks with very rapidly occurring stimuli. With successive tasks, the signal and non-signal are not present at the same time and therefore the signal is distinguished from a non-signal by referring to a representation in memory (Matthews, Davies, & Holley 1993). By contrast, simultaneous discrimination tasks contain the signal and non-signal at the same time, and these can be compared directly with each other. Successive tasks require more mental effort, attentional capacity and working memory (Matthews et al., 1993).

The sensitivity decrement is associated with a genuine loss in the observer's ability to discriminate signals from non-signals and can be substantial (See et al., 1995). Signals that are easy for alert observers to detect at the beginning of the vigilance task may become fairly difficult later in the task. This suggests that operators' ability to perceive some significant events may deteriorate over time. Factors that do not appear to mediate sensitivity decrements include stimulus duration, signal amplitude, sensory modality, signal regularity, event asynchrony, spatial uncertainty, signal probability, training time, feedback regarding training results, feedback on detection rates during vigil, and the length of the vigil (See et al., 1995). This suggests that interventions commonly aimed at improving vigilance performance, such as reducing shift lengths, training and providing feedback, are unlikely to be effective in reducing the sensitivity decrement (although they may assist other aspects of vigilance performance).

The response bias or criterion is the second process involved in detection. It refers to the operator's decision process regarding the presence of a signal and is defined as 'the tendency to favour one response over the other' (MacMillan & Creelman, 1991, p. 31). Responses to signals occur when a combination of encoded information (sensitivity) and the operator's objectives and task knowledge exceed a detection or response criterion (Balakrishnan, 1998).

The response criterion is under the operator's control. While some operators may be conservative in their detection decisions, others may be more likely to take risks. Those who are more 'risky' are likely to detect more significant events but are also likely to have high FA rates. On the other hand, operators who are conservative have few FAs but also miss many significant events (Davies & Parasuraman, 1982).

The consequences of hits, FAs and misses influence the appropriateness of risky versus conservative decisions (Wickens & Hollands, 2000). For instance, when there are dire consequences for missing a significant event, it may be better to decide that a significant event is present and risk a FA than to 'miss' it. With CCTV, the potential consequences vary according to the context and significant event but are likely to influence the response criterion. The response criterion's optimal value is influenced by signal probability, payoffs, and human performance (Wickens & Hollands, 2000).

Based on expectancy theory, over time observers alter their tendency to decide whether a target is present or not (MacMillan & Creelman, 1991). Initially observers base the response criterion on subjective expectations of how often the signal occurs. If a signal is missed, the expectation is developed that there will be fewer signals than originally expected and the response criterion is altered. This adjustment in turn increases the chances of missing the target, and a vicious cycle develops (Parasuraman, 1984). In an experiment where target rates decreased from 50 to 1 percent, misses increased from 7 to 30 percent (Wolfe, Horowitz, & Kenner, 2005). When signals are infrequent, the response criterion is set higher.

The prevalence of signals is linked to the response criterion and detection performance. When signals occur infrequently, the number of misses increases and FAs decrease, even when observers have the opportunity to correct response mistakes and perform the search again (van Wert, Horowitz, & Wolfe, 2009). This has serious implications for visual search tasks with important social implications, such as airport baggage screening, tumour detection, and CCTV surveillance. With these tasks, signals occur infrequently and an altered response criterion is likely to result in increased misses with their associated consequences (van Wert et al., 2009). The prevalence effect implies that targets that occur frequently may be detected, but when the same targets occur seldom they may be missed.

The expectancy account has received some support, but various studies have indicated that vigilance decrements occur even when expectations are managed, for example through training (Milosevic, 1975; Parasuraman, 1976, 1979; Parasuraman & Davies, 1976; Williges, 1973). However, expectancy theory has been used successfully in the development of TIP for screeners where the frequency of threat objects is manipulated (Hancock & Hart, 2002). SDT makes a valuable contribution regarding certain factors involved in deciding whether a signal is present or not. However, it tends to focus on expectancies and tendencies to respond and does not go far enough in explaining why particular stimuli are recognised as being significant and others are not. Although it acknowledges the importance of factors such as noise and detection thresholds, it ignores various processes that contribute to recognising the significance of perceived stimuli. These include, for example, SA and its associated mental models, schemas, and attention sets. It says little about how sensory information is interpreted, and subsequent decisions regarding the presence of significant events. In some instances, operators are likely to look at the appropriate area of the display and see certain target behaviours (TBs), but may not recognise their significance. This is not related to expectancies or the response criterion, but to other cognitive processes. Therefore SDT only provides a partial explanation of decisions in vigilance tasks.

Despite the frequent application of SDT to vigilance research, its suitability has been debated and at times it has been applied inappropriately and uncritically (Long & Waag, 1981). Conceptual and statistical issues form the core of the debate. A key issue focuses on assumptions regarding normality and the use of parametric analyses (See et al., 1995). The application of SDT to group rather than individual data, as originally intended, has also been criticised (Craig, 1977). The reliability of data is often low (Davies & Parasuraman, 1982). This is related to low signal frequencies and consequent low response rates. Much SDT research also tends to use supra-threshold signals which are easy to detect, resulting in near perfect detection rates and low FA rates (Long & Waag, 1981).

Lastly, various factors that could affect vigilance performance are excluded from the SDT framework, including for example, boredom and fatigue, lapses in attention, probability matching and learning (Long & Waag, 1981), overall vigilance levels, search patterns, and memory processes (Davies & Parasuraman, 1982). Like other theories used to explain vigilance performance, the SDT framework provides an incomplete account of vigilance performance. Despite these shortfalls, SDT provides the basis for understanding certain key processes underlying vigilance behaviour (Davies & Parasuraman, 1982). The distinction between sensitivity and the response criterion and the link with expectancies are particularly relevant to the current research. Although they are not measured in this study, they contribute to an understanding of the viewing and detection process.

2.2.3. Multiple resource attention theory.

Multiple resource theories of attention explain the vigilance decrement in terms of the depletion of attention resources (Wickens & McCarley, 2008). These theories propose that attention is not a unitary, undifferentiated pool of general resources (e.g., Polson & Friedman, 1988; Boles, 2002; Boles & Law, 1998; Wickens, 2002) but consists of numerous different resources (Warm et al., 2008). The resources are seen as being structurally and neurologically separate so that different resources can be allocated to tasks requiring different resources, such as auditory or visual processing (Wickens, 2008). Resources include, for example, stages (perceptual and cognitive versus choosing and executing a response), sensory modalities (auditory versus visual), codes (visual versus spatial) and channels of visual information (focal versus ambient) (Wickens, 2002).

Multiple resource theories contend that the vigilance decrement is not dependent on time, but on the continuous demand for and depletion of resources (Smit et al., 2004). The viewing of CCTV displays is largely perceptual and cognitive, visual (focal and ambient) and uses visual and spatial codes. These resources are depleted most during visual monitoring. The addition of other tasks to the operator's job could place demands on the same resources and deplete resources further. Alternatively, additional tasks may require different resources from monitoring and therefore are unlikely to affect detection levels unless they remove the gaze from displays.

Research has examined the use of resources when more than one task is performed simultaneously. Greater interference is predicted between tasks that draw on the same resources than between those that use different resources (Wickens, 2002). In the CCTV context, verbal responses to a low priority alarm should not interfere with monitoring activities visually on camera. However, searching multiple CCTV displays or writing reports could deplete visual resources and influence detection.

Effort in itself demands resources and this is clear in vigilance situations where:

> '… the mental demands of simply watching and waiting for rare events to occur have been well documented (Warm, Dember, & Hancock, 1986), even

under conditions in which performance does not suffer a great deal.' (Wickens & McCarley, 2008, p. 120).

Task difficulty modulates the amount of effort and resources required, with highly automatic tasks requiring less mental workload and fewer resources, making them less vulnerable to performance decrements than tasks that make high demands on working memory (Wickens & McCarley, 2008). With CCTV, certain significant events are easier to detect than others, and are therefore likely to make different demands on resources. Resource theory provides a useful way of identifying the information processing requirements of tasks and contributes to an understanding of the types of tasks that can be effectively combined in jobs.

2.2.4. Mindlessness theory.

The mindlessness model of vigilance proposes that when tasks are monotonous and responses are required very infrequently, attention is withdrawn from the task which is then performed in a routinised and thoughtless manner (Robertson, Manly, Andrade, Baddeley, & Yiend, 1997). This mindlessness is seen as an endogenous loss of potency of the supervisory attentional control system, and distinct from decreased arousal (Dickman, 2002). Multiple resource theories acknowledge the existence of attention failures, but attribute these to depleted cognitive resources rather than mindlessness (MacLean et al., 2009).

Boredom is incorporated under the umbrella of the mindlessness theory of vigilance (Helton, Weil, Middlemiss, & Sawers, 2010). Theorists in support of the mindlessness theory argue that the boring and repetitive nature of vigilance tasks leads to the withdrawal of conscious attention (Smallwood, Davies, Heim, Finnigan, Sudberry, O'Connor, et al., 2004). Attention plays a central role in most conceptions of boredom which is seen as an inability to engage with the task and to sustain attention despite being free to do so (Carriére, Cheyne, & Smilek, 2008). Three types of boredom are distinguished (Cheyne, Carriére, & Smilek, 2006). These are frustration of the ability to engage attention, the forced engagement of attention and an inability to maintain attention. It is possible that all of these could apply in the CCTV context. Boredom based on specific situations is distinguished from boredom proneness, which emerges from a lack of motivation and effort (Cheyne et al., 2006). Thus the concepts of boredom, motivation, interest and effort are linked (Carriére et al., 2008). Boredom is related to various types of attention failures and human errors which can be costly (Borell-Carrió, & Epstein, 2004; Robertson, 2003). However, such attention failures are not only attributed to boredom and mindlessness and are discussed in more detail later in this chapter.

Support for the mindlessness theory is based largely on the Sustained Attention to Response Task (SART) (Helton et al., 2010). This typically uses simple, meaningless stimuli and requires participants to respond to non-signals and withhold responses to signals (Cheyne, Solman, Carriére, & Smilek, 2009). This is very different from vigilance tasks in real world contexts. The SART has been criticised by a number of researchers on the grounds that it assesses impulsiveness (Helton et al., 2010) and response strategy (Helton, 2009) rather than sustained attention.

The mindlessness model has also been challenged on theoretical grounds. Given the high mental workload imposed by vigilance tasks, several studies concluded that poor performance is the result of limitations in information processing resources rather than an approach to the task that is lacking in thought (e.g., Grier, Warm, Dember, Matthews, Galinsky, Szalma, & Parasuraman, 2003; Helton, Hollander, Warm, & Matthews, Dember, Wallart, et al., 2005; Johnson & Proctor, 2004; Matthews, Davies, Westerman, & Stammers, 2000). It is fairly well established that the high workload of vigilance tasks stems from information processing demands inherent in the tasks rather than efforts to overcome boredom (Helton et al., 2005; Hitchcock, Dember, Warm, Moroney, & See, 1999). This implies that vigilance jobs should be designed to reduce mental workload and stress.

For operators, reductions in mental workload could be achieved through shorter shifts, observing fewer CCTV displays, performing fewer additional tasks while observing displays, and providing training aimed at increasing operators' understanding of how significant events unfold and their recognition of suspicious events that tend to surround incidents. However, the effectiveness of these interventions needs to be evaluated. The mindlessness theory has implications for how actively operators monitor cameras as passive monitoring is less likely to result in the detection of less visible events. For the purpose of this research, passive monitoring is seen as being similar to a mindless approach to surveillance whereby attention is withdrawn from the task and little effort is made to detect significant events. The applicability of the mindlessness theory to CCTV needs to be evaluated in future research.

The shortfalls in theories that focus on the vigilance decrement have prompted a recent change in the direction of vigilance research from identifying factors that impact on the vigilance decrement, to how attention processes and resources are deployed during vigilance (Warm et al., 2008). Although these processes also apply to contexts that do not require vigilance, they offer useful insights into vigilance intensive jobs. Task dis/engagement has been examined in the vigilance context although strictly speaking, it is not a vigilance theory. It is included in this section because it is a similar process to mindlessness, it provides an explanation of attention deployment over time that is not dependent on the existence of a vigilance decrement, and it offers an alternative perspective on attention processes that occur during task performance.

Task disengagement is defined as ‘A state of reduced allocation of attentional resources to environmental task-related stimuli’ (Cheyne et al., 2009, p. 98). This is similar to mindlessness but is applied to a variety of tasks including those that do not require vigilance. During task disengagement, attention resources are transferred to stimuli and cognitions that are unrelated to the immediate task, or off-line processing. This involves various types of mind-wandering and cognitions. These cognitions can have either positive or negative effects on learning and performance (Kanfer & Ackerman, 1996). However, various studies have found that mind-wandering reduces the ability to perceive and respond quickly to external events (He, Becic, Lee, & McCarley, 2011). Some of the cognitions that occur during mind-wandering include content that is not related to the task (Smallwood et al., 2004). However, cognitions about task performance also occur, especially when there are concerns about successful performance (Gilden & Hancock, 2007). Mind-wandering therefore includes both task unrelated thoughts (TUTs) (Smallwood, Fishman, & Schooler, 2007) and task related interference (Sarason, Sarason, Keefe, Hayes, & Shearin, 1986). Task disengagement and mind-wandering are associated with attention failures and errors, some of which have serious consequences (Dockree, Kelly, Robertson, Reilly, & Foxe, 2005).

Disengagement results in lapses in the maintenance of task goals and even momentary lapses can be highly problematic. For example, extremely brief lapses account for a large proportion of road traffic accidents.

Cheyne et al. (2009) found that task disengagement is most likely to occur when there are long, inactive periods between significant events. Their findings also suggest that engagement is unstable and constantly fluctuates. These authors propose a three-state model of attention disengagement. The first state involves 'tuning out' (p. 99) and decreased awareness of moment-to-moment changes in environmental stimuli. However, the person remains aware of the general task environment, divides attention between the task and cognitions that are not related to the immediate task, and still devotes sufficient resources to the task to prevent major errors.

The second state of disengagement involves mind-wandering in a less conscious manner. There is decreased awareness of generic task-relevant stimuli in the environment, and a more automated manner of responding. This is similar to the concept of 'zoning out' (Smallwood & Schooler, 2006). Responses to near misses decrease, but blatant errors are likely to re-deploy attention to the task (Cheyne et al., 2009). The third and deepest level of disengagement involves little awareness of task demands and the person only responds to highly intrusive task events. This state is accompanied by behavioural indicators of mind-wandering. Motor restlessness is associated with fatigue and has been found to increase significantly over a 50-minute watch (Galinsky, Rosa, Warm, & Dember, 1993).

The task (dis)engagement paradigm contributes to an understanding of individual variations in patterns of vigilance performance over time without relying on a steady vigilance decrement. It allows for engagement and disengagement to alternate with varying patterns. It also suggests that once disengagement from task goals has occurred, re-activation of goals is important (Posner & Petersen, 1990). However, little is currently known about the mechanisms involved re-engagement. It is possible that IGOs could assist with the re-engagement process.

2.3. Implications of Theories for Vigilance Performance

Vigilance theories make different predictions for detection performance. Arousal theory predicts that vigilance tasks with a greater mental workload and complexity will be associated with a smaller performance decrement because arousal levels will be maintained better through decreased monotony and greater stimulation (Smit, Eling, & Coenen, 2004). The mindlessness theory makes a similar prediction that by increasing task challenge and mental workload, the vigilance decrement will be reduced (Helton et al., 2010). On the other hand, multiple resource theory predicts that prolonged tasks with a greater workloads lead to greater resource depletion and poorer performance (Smit et al., 2004). While there is mixed support for arousal theory's prediction, several studies support multiple resource theory (Craig, 1987; Matthews & Davies, 2001; Parasuraman & Mouloua, 1987; Parasuraman, de Visser, Clarke, McGarry, Hussey, Shaw, & Thompson, 2009).

SDT predicts that sensitivity decreases over time, especially for tasks with successive discrimination where visual stimuli are compared with events in memory (Matthews, Davies, & Holley 1993), as with CCTV. Similarly, the response criterion is adjusted over time resulting in different detection rates (MacMillan & Creelman, 1991). Factors such as signal frequency and the implications of signals influence the response criterion (Wickens & Hollands, 2000) which can become stricter or more lenient.

With the exception of task disengagement theory, all the above theories focus on the vigilance decrement rather than the overall level of vigilance. They appear to be based on the assumption that a vigilance decrement will occur. Indeed, the majority of vigilance research has found empirical evidence of a vigilance decrement. However, most of these studies are based on laboratory research using simple, synthetic and meaningless stimuli (Craig, 1984) and students as participants (Childs, 1976; Cunningham, Scerbo, & Freeman, 2000; Goldstein, Johnston, & Howell, 1969; Helton, Dember, Warm, & Matthews, 1999; Helton et al., 2007; Hollenbeck et al., 1995; Jerison, 1959; Rose, Murphy, Byard, & Nikzad, 2002; Sawin & Scerbo, 1993, 1994; Wiener, 1973). Comparatively little research has used real world tasks with job incumbents as participants. In addition, the existence of a vigilance decrement is likely to vary across types of task (e.g., air traffic controllers versus CCTV surveillance operators) and operations as it is probably related to numerous factors such as motivation, shift length and rest breaks, job design, control room culture and amongst other factors (Donald, F., 2008).

Vigilance theories have seldom acknowledged that observers are not necessarily passive and that they have a certain degree of volition in self-regulating their attention and controlling it in an effortful manner. Self-regulation refers to the ability to monitor the extent to which one's attention is focused on the task and to regulate task unrelated thoughts and task related interference (Smallwood et al., 2004). Self-regulation has been found to have a positive effect on reducing cognitive interference, thereby improving motivation and performance (Kanfer & Ackerman, 1996). Kanfer and Ackerman (1996) describe two major types of self-regulatory skills - emotion and motivation control. The self-regulation of emotion reduces negative emotions that affect performance, while motivation control is aimed at ensuring that attention and effort are focused on the task, despite emotions and attitudes towards the task. The distinction between these is useful as it links emotions (including boredom), attitudes, and cognitions about the task and unrelated matters to motivation, effort and performance. This framework has the potential to explain many of the dynamics that occur during vigilance performance but requires future validation in operational vigilance contexts.

The degree of volitional attention regulation that occurs is likely to vary with across vigilance tasks and types of samples. This is consistent with the finding that vigilance performance and task engagement vary across tasks (Smallwood, Obonsawin, & Heim, 2003). It is likely that participant's motivation and related efforts to maintain task engagement over time are related to the subjective meaningfulness of the task and potential consequences of poor performance. These factors are likely to differ between laboratory experiments using simple, meaningless tasks and research using tasks that are more similar to those encountered in real life. In addition, the meaning of these tasks is likely to differ for participants who perform the tasks as part of their everyday jobs, compared with participants for whom the tasks are novel. Future research is needed that examines the dynamics regarding volitional attention self-regulations and vigilance performance with complex tasks that are similar to those performed in real life.

2.4. Contributions of Vigilance Research to an Understanding of CCTV Operator Performance

Vigilance research has made a valuable contribution by highlighting numerous factors that affect vigilance performance, including aspects of the signal, task, environment and operator. These are summarised in Table 2 which is based on Mackie (1987), Balakrishnan (1998), Koelega (1996) and See et al. (1995). These factors apply to the CCTV environment and should be considered in designing interventions.

Table 2. Factors affecting performance in vigilance tasks (Adapted from Mackie, 1987)

Signal characteristics	**Task characteristics**	**Individual differences**
Signal rate	Duration	Motivation (intrinsic, extrinsic)
Signal complexity	Number of sessions	Mood, morale
Signal regularity	Monitoring load, secondary tasks	Work/rest patterns
Signal probability	FA rate	Circadian rhythms
Signal conspicuity	Irrelevant stimulation	Sleep loss and quality
Signal uncertainty	Payoffs, rewards	Biofeedback
Background event rate	Technology, e.g., automation	Operator pacing
Sensory mode(s)	Knowledge of results	Illness, injury
Discrimination requirement (simultaneous vs successive, sensory vs cognitive)	Environmental stressors (noise, heat, cold, vibration, carbon monoxide, hypoxia)	Drug use Smoking Exercise, physical work
Critical features	Confidence judgments	
Inter-signal interval	Encoding	

The second key contribution is that vigilance has come to be seen as a multidimensional construct, consisting of many qualitatively different mechanisms that cannot be explained in terms of a single mechanism (Koelega, 1996). No single theory explains all aspects of vigilance performance and therefore no single theory can be applied to the current study. A synthesised theory that integrates the mechanisms is unlikely to be developed due to the large range of variables that affect vigilance performance and the fact that it is context specific. However, by understanding the multidimensional nature of vigilance, various avenues for interventions are created. For instance, learning theories indicate the need for rest breaks to address fatigue. Arousal theory suggests that mental workload should be assessed and arousal maintained over time. SDT focuses on sensitivity and the decisions made by operators. It emphasises the importance of manipulating expectancies through interventions such as training and providing feedback. While multiple resource theory also examines workload, it suggests that the demand for particular resources be used as the point of departure for interventions.

Despite the contribution of vigilance studies, their value has been debated (e.g., Adams, 1987; Koelega, 1996; Mackie, 1987; Parasuraman, 1987; Pigeau et al., 1995; Sawin & Scerbo, 1995; Wiener, 1987). Much of the debate has focused on the tendency to conduct vigilance research in laboratory settings rather than operational environments (Craig, 1984). While some researchers have cited examples of the successful application of findings (e.g., Parasuraman et al., 1987), others have questioned their relevance and predictive value in real world environments (Adams, 1987; Koelega, 1996; Mackie, 1987; Pigeau at al., 1995; Wiener, 1987). A major reason for this is the difficulty in generalising results of laboratory experiments to real world operating environments. This is based on differences between vigilance tasks used in laboratory experiments and those faced by operators in work contexts. Laboratory experiments tend to use tasks that are extremely simple compared with the complex and often dynamic tasks performed in operational settings. In addition, the synthetic stimuli that are frequently used in laboratory experiments lack inherent meaning and this is likely to affect the motivation of participants.

The nature of the task and displays differ between simple laboratory experiments and operational contexts. In terms of displays, operators in the real world context typically observe anything from 3 to 30 camera scenes, yielding multiple sources of data (Donald, F, 2008). Given numerous displays, mental models of the interrelationships between different cameras and scenes are needed to anticipate and track movement. An understanding of the meaning of actions depicted in CCTV footage, divided and selective attention, effective scan patterns and a large working memory capacity are needed to operate the cameras and process numerous sources of data. Display backgrounds in operational settings tend to be cluttered and contain a variety of types of images. Display complexity is increased by the fact that displays are two-dimensional representations of three-dimensional scenes and events and objects may be hidden or partially obscured. This is particularly relevant to situations where people deliberately obscure objects and situations from the camera's view to avoid detection. Consequently abstraction and prediction are required in interpreting scenes. The displays monitored by operators are dynamic and typically contain one or more moving objects, representing a continuous event rate. This contrasts with much vigilance research which uses discrete events. In addition, several objects could change at the same time, at different rates, and in different directions (Donald, F., 2008). Continuous display changes require the on-going updating of SA (Endsley 1995), introducing additional information processing demands that are not present in traditional vigilance research (Donald, F., 2008).

The nature of significant events in CCTV also places additional demands on operators. Significant events include, for example, the presence or behavior of people, changes in process, anomalies in expected conditions, verification of standard operating conditions or protocols, or the detection of specific threats or circumstances (Aldridge, 1994). These events may involve small, covert, transient actions that take merely a moment to execute, recognising patterns or trends in behaviour, or identifying intermittent events that are embedded in cluttered, changing backgrounds of displays (Donald, F., 2008). Thus the perceptual salience of relevant data varies and the significant events themselves consist of complex images.

Recognising complex situations that occur in real life environments requires consideration of a range and diversity of elements and possible combinations of significant events, conditions and characteristics (Mackie, Wylie, & Smith, 1994). This applies to CCTV where successful operators use non-verbal cues and patterns to comprehend actions and predict significant events in a proactive manner. This underlines the importance of SA in CCTV surveillance. Based on Endsley's (2000) levels of SA, operators require high levels of SA whereby they interpret data, match patterns, create meaning and anticipate future events.

This contrasts with traditional vigilance research where little knowledge other than the form of the significant event is required for detection and the emphasis is on perceiving signals at a sensory level, or Endsley's lowest level of SA.

In the current research, previously described measures are taken to reduce the level of SA required of participants. However, in real life operations the ability to maintain up-to-date SA and remain engaged with the surveillance of displays is challenged by job design. In addition to controlling and monitoring the CCTV surveillance system, operators typically perform a range of other tasks (Donald, C., 2001). These include, for example, responding to situations and significant events they observe, liaison with other parties and services (e.g., emergency services, police, traffic authorities), accessing various data bases (e.g., vehicle registrations), following up on significant events, and writing reports. Depending on how tasks are structured, these activities may distract from the monitoring and detection task. Unlike experimental participants, operators are part of broader systems and perform other tasks that add to workload and job complexity (Donald, F., 2008).

CCTV surveillance draws on complex data with multiple inter-related facets. This suggests that the traditional approach of studying vigilance by isolating and manipulating simple variables is likely to yield different results from research on real world tasks (Donald, F., 2008). Because CCTV detection involves both perceptual and cognitive processes based on inter-related aspects of the display, the composite impact of the overall display is more important than the effects of discrete characteristics or stimuli themselves.

The second drawback of traditional vigilance research is that it reveals relatively little about what is occurring during the vigilance task (Koelega, 1996, p. 280):

> '… there are serious difficulties in the application of *all* the models. A failure to detect and respond may occur because of sleepiness, *or* by decreases in sensitivity (d'), *or* by failures to observe, *or* by habituation to a level below the criterion, *or* by an increase in criterion based on expectancies, *or* may result from the interaction of these various processes, possibly with other processes as well.'

This comment is likely to apply to operators in the CCTV context although no empirical research has been found that examines it. In addition, vigilance theories do not contribute to an understanding of operators' strategies for searching displays, how these influence alertness and detection performance or whether and how operators regulate their attention and task engagement. They do not provide a microanalysis of the attention processes that occur throughout the shift and how these affect performance. However, task dis/engagement theory and its moment-by-moment analysis of the patterns of task engagement make a significant contribution to an understanding of these processes.

Search strategy is one of the processes that occurs during the vigilance task that has not been included in vigilance theories. It influences where observers direct their gaze and what they deploy attention to. With CCTV, many significant events are highly transitory in nature, and are missed if focal attention is not directed at the relevant area at the right time. This contrasts with vigilance tasks in some other contexts, such as baggage screeners where X-rays depict static objects and the same image is present for several seconds. Further, many perpetrators attempt to hide behaviours from the CCTV camera's view or disguise them, making detection more difficult. Therefore search strategy is extremely important and is likely to have a strong influence on overall detection rates. Despite this, the search strategies and eye movements of people working in CCTV have been neglected by researchers. Such research has been conducted in other contexts (Castelhano & Henderson, 2008; Henderson, Brockmole, Castelhano, & Mack, 2007; Underwood & Foulsham, 2006). Little is known about how operators search through CCTV displays except that strategies vary from random searches, to relying on 'gut feel' to identify suspicious events, to more systematic search by camera number or geographical area (Wells et al., 2006). Some operators limit proactive monitoring to particular areas or times when significant events are most likely to occur. However, these limited findings do not stem from vigilance research with a focus on vigilance.

Having provided an overview of vigilance research, the background to the intervention and its rationale are explained in the next chapter.

CHAPTER 3: BACKGROUND TO THE INTERVENTION

Operators performing proactive CCTV surveillance typically perform vigilance intensive work for extended periods of time (Donald, 2001). As with other vigilance-intensive tasks, performance may be sub-optimal in some CCTV contexts. Various researchers have commented that operator performance in CCTV is frequently inadequate (Cameron et al., 2008; Carli, 2008; Donald & Andrew, 2003; Donald et al., 2007; Gill & Spriggs, 2005; Keval & Sasse, 2006; Squires, 2010; Wells et al., 2006), but the levels of detection required to be considered ‘optimal’ or ‘adequate’ have not been specified. For certain significant events with serious consequences (e.g., the detection of bombs) a one hundred percent detection rate is needed, whereas lower rates could be acceptable in other contexts. Detection rates are likely to vary across CCTV systems, organisations and industries and may be satisfactory in some situations. A few studies have been conducted on detection rates for CCTV surveillance and have obtained a wide range of results. The highest detection rate found (91%) to date was obtained in a study where simulated city centre incidents had to be identified, such as theft of a handbag (van Voorthuijsen, van Hoof, Roubik, Klima, & Bernas, 2005). A study requiring participants to detect a person carrying a visual conspicuous open umbrella in scenes where no other umbrellas were present, reported a, 85% detection rate (Wallace, Diffley, Baines, & Aldridge, n.d.). These contrast with research using footage from diamond processing plants which obtained a 25% detection rate (Andrew et al., 2003). However, the comment has been made that ‘there is a need for a test that would be able to be routinely used in a CCTV control room throughout the course of a normal shift to provide management with operational performance data’ (Neil et al., 2007, p. 1). While the intervention developed in the current research could contribute to meeting this need, it aims to go beyond providing performance data and to facilitate the detection of significant events.

Researchers typically evaluate CCTV effectiveness using measures such as reduced crime rates and public perceptions of crime and safety (e.g., Gill & Spriggs, 2005; Wells et al., 2006). These evaluations are useful in providing overall indicators of the systems’ aims, but provide little information on statistics regarding the number of hits, misses and false alarms

(FAs) made by operators. Moreover, statistics on operator performance are unlikely to be accurate due to measurement issues (Donald, 2001). If a significant event is not detected, its presence may remain unknown and cannot be measured as a 'miss' (Donald, C., 2008). Although some organisations calculate statistics on operator detection rates, they are unlikely to be released into the public domain as they are considered confidential.

In many contexts where CCTV is used, significant events occur infrequently and are not predictable (Donald & Donald, 2008). Weeks or even months may pass without a significant event occurring. Performance evaluation during this time is therefore based on 'non-events.' In addition, it is not always apparent when significant events have occurred. Some significant events are only detected when tapes are reviewed at a later stage rather than in real time (Donald, 2001). Reviewing tapes is a time consuming process that cannot be undertaken for all footage. Therefore evaluations are likely to be based on small samples of tapes and some significant events will inevitably be missed. Further, it is difficult to evaluate the extent to which operators deploy attention to displays, especially when there is little activity to be observed (Donald, C., 2008).

Despite challenges regarding the evaluation of proactive CCTV surveillance, research suggests that there is space for improvement in operator performance. Some evidence of the vigilance decrement has been found in an experimental setting involving a realistic task where operators were participants (Andrew, Landgrebe, & Donald, 2003). In addition, operators tend to engage in activities that remove their attention from displays and this could affect detection levels (Smith, 2004). The vigilance decrement is likely to be found in situations where there is high mental workload (Temple et al., 2000; Warm et al., 2008), inconspicuous significant events (Becker, Warm, Dember, & Hancock, 1991), temporal and spatial uncertainty, changing displays (Parasuraman & Mouloua, 1996), and successive discrimination tasks (Matthews, Davies, & Holley 1993). In successive tasks the signal and non-signal are not present at the same time and therefore the signal is distinguished from a non-signal through reference to a representation in the observer's memory (Matthews et al., 1993). These conditions apply to CCTV surveillance, suggesting that the vigilance decrement is likely to occur in this context.

Vigilance, monitoring and detection processes are complex and make high demands on mental workload (Parasuraman & Mouloua, 1996). The nature of significant events and the CCTV system (e.g., ratio of displays to operators), tasks performed in addition to surveillance, the overall level of vigilance, the vigilance decrement (Parasuraman, 1984), motivation, habituation (Broadbent, 1952), inhibition (McCormack, 1959), arousal (Matthews & Davies, 1998), expectancies, sensitivity, the response criterion (MacMillan & Creelman, 1991), signal characteristics (Mackie, 1987), and SA (Endsley, 1995) all contribute to detection performance. Successful detection results from the complex interaction of these processes.

It is not known how successful operators are at their detection task or whether they are vulnerable to the vigilance decrement. With the exception of the Andrew et al. (2003) study, actual detection rates and vigilance decrements for operators have not been established. The current study examines the detection rates achieved and potential vigilance decrements in a video surveillance task based on diamond processing plants. The numerous processes contributing to vigilance performance indicate that no single intervention is likely to address all performance issues (Donald & Donald, 2008). However, an intervention aimed at maintaining vigilance over time and enhancing the detection of significant events would still be practically beneficial. One of the aims of the current research is to contribute to the development of such an intervention.

3.1. Basis for the IGO Intervention

The intervention developed in the current research builds on an initial study by Andrew et al. (2003). It combines two key aspects – the use of graphics and the insertion of objects. Both aspects have been applied successfully in different contexts. Graphics have been applied to icons, signs, symbols, and pictograms used in computer or graphical user interfaces, labels and information leaflets on medicines, chemicals, household appliances, safety and warning systems, and transport systems (McDougall, Tyrer, & Folkard, 2006).

Images are used to represent objects or functions and convey information (Thatcher, Mahlangu, & Zimmerman, 2006), develop and reinforce knowledge, change behaviour (Dowse & Ehlers, 2004; Miller & Stanney, 1997; Yin, Dreyer, van Schaick, Foltin, Dinglas, &

Mendelsohn, 2008), and activate computer or machine functions (Wogalter, Sojourner, & Brelsford, 1997). IGOs could share some of these functions, such as developing and reinforcing an understanding of what comprises significant events and how they unfold, and changing behaviour or the manner in which operators search displays. In addition, they could assist the recognition of significant events by modeling or representing key aspects of significant events.

Inserted objects are used in threat image projection (TIP) for baggage X-ray screening in aviation (Neiderman & Fobes, 2005) and artificial signal detection (ASI) mainly for inspection in manufacturing (Wilkinson, 1964). Although the CCTV context is different in some respects from X-ray screening, and IGOs serve somewhat different purposes from icons and pictograms, these contexts provide indicators regarding IGO design and how IGOs could potentially be used to enhance the recognition of significant events or incidents.

The use of graphics is assumed to have a number of advantages over text, such as conveying large amounts of information efficiently and concisely in a small space, being accessible across language barriers and cultures, and to people with low literacy levels (Isherwood, McDougall, & Curry, 2007). They are more rapidly perceived than text and more easily seen under degraded conditions (Hancock, Rogers, Schroeder, & Fisk, 2004). Consequently they provide large amounts of information at a glance (Laughery, 2006). There is evidence that icons and pictograms can be learnt more easily (Wiedenbeck, 1999) and remembered better than text (Blankenberg & Hahn, 1991; Mansoor & Dowse, 2007).

Dual coding theory (Paivio, 1975; Johnson, Paivio, & Clark, 1996) assists in explaining the cognitive processing advantages of graphics or images (Wogalter et al., 1997). The theory proposes that visual and language or text-based stimuli are coded separately in the brain. Images tend to be coded in both systems, resulting in richer memory representations and mental models (Thatcher at al., 2006). Multiple resource theories of attention support the efficient processing of images by suggesting that they are processed during early visual coding. This allows resources to be used for higher-order processing, such as problem-solving. It has also been suggested that images reduce mental workload by reducing memory recall processes (Rogers, 1989, cited in Thatcher et al., 2006). As IGOs are images, the same benefits and processes are seen as applying to them.

TIP has proved an effective means to heighten screener performance (Bennett & Fobes, 1994). It refers to the electronic projection of images of objects that pose a threat, such as weapons and explosives, into real time X-ray images of actual bags at airports (Neiderman & Fobes, 2005). Screeners are required to detect these projected images. TIP has been used successfully with X-ray baggage screeners (referred to as 'screeners') at airports (Foulkes et al., 1997) and in simulated studies with professional sonar operators (Mackie et al., 1994). In these settings, TIP images have also been used for performance measurement and feedback, suggesting that IGOs could be used for this as well.

TIP is based on ASI research conducted in the 1960s (e.g., Baker, 1960; Wilkinson, 1964). Faulty items were placed amongst other items for inspection in order to measure how many faults were noticed or missed and how many FAs occurred. Early ASI studies obtained limited support (Baker, 1960; Garvey, Taylor, & Newlin, 1959; Wallis & Newton, 1957; Wilkinson, 1964). However, a more recent study using sonar signals found improved detection rates when feedback was provided (Mackie et al., 1994). Feedback is an important component in ASI effectiveness and could also play a role in the CCTV context. However, an examination of the effects of feedback is beyond the scope of the current study. Inserted objects were also found to alter the response criterion and perceptual sensitivity (Mackie et al., 1994).

Similar visual analysis processes seem to underpin the jobs of operators and X-ray screeners, but the nature of the detection task differs (Donald & Donald, 2008). Both positions require job incumbents to conduct proactive surveillance in real time. This involves forming mental pictures of what is happening and drawing upon SA in the process. However, CCTV displays have a number of differences from X-rays and the ASI contexts. CCTV displays contain a number of factors that are not constant due to the ability to zoom in, different camera views, and most importantly, the dynamic nature of displays where people and objects in the scene are moving. Compared with TIP and ASI, this poses additional challenges regarding IGO design and placement. In addition, the significant events themselves are at times complex as they do not consist merely of objects, but include behaviours and patterns of events as well. Certain significant events take a while for perpetrators to set up and implement, but others occur in a split second, posing additional challenges and choices for IGO design.

Operators who are more successful at detecting either of these types of significant events are likely to use accurate mental models and SA to assist in recognising the behaviours that signify significant events (Donald, F., 2008). An understanding of behaviours assists in anticipating events and recognising them as they unfold (Donald, 2004). Therefore, operators need to observe objects, behaviours and patterns, while screeners focus only on objects. The range of behaviours and events that suggest a significant event is taking place implies that IGOs could take on a variety of forms to reflect these behaviours and events. Therefore careful consideration needs to be given to which IGO characteristics are likely to be most effective in maintaining vigilance and enhancing detection.

The X-ray context provides a clear indication of the type of images to use as TIP objects, such as weapons or parts of weapons. For CCTV, it is not as clear which types of IGOs are appropriate (Donald & Donald, 2008). In view of the complexity of CCTV displays and the vast array of scenes recorded, a wide range of images could potentially be used. These could include, for example, simple shapes, items related to significant events (such as valuable objects or weapons), symbols, humourous items, and clips of real recordings of significant events. Some of these may be relevant to significant events, while others may be arbitrary and bear no relation to the types of events operators are required to detect.

Two existing studies on IGOs have used a mixture of IGOs with a range of characteristics. The first study used arbitrary IGOs that were not related to the detection task, such as dogs and numbers (Neil et al., 2007). The emphasis was on the detection of IGOs when different monitor configurations were used. Neither the influence of IGOs on the detection of actual significant events nor the detection of significant events themselves was evaluated. Because the current research focuses on how IGOs could assist the recognition of significant events, this study has little relevance to the current research.

The second study used a mixture of IGOs with different characteristics varying from simple shapes to static figures of persons, to static figures in postures associated with theft behaviour (Andrew et al., 2003). Some IGOs represented objects or behaviours associated with significant events, while others were not related to significant events. The IGOs varied in shape, size, type and location.

The Andrew et al. (2003) study used a realistic surveillance task in the diamond industry and aimed to establish whether IGOs could enhance detection rates (Donald, Andrew, & Landgrebe, 2007). The duration of the test session was three hours and 49 minutes with three short breaks that were evenly spaced. After the first 70-minute interval, the treatment (received the IGOs) and control groups (received no IGOs) obtained similar scores (Andrew et al., 2003). Thereafter performance began to diverge, with the treatment group achieving a slightly increased (but not statistically significant) detection score after the second interval, and a decreased detection score at the end of the third 70-minute interval. This suggests that the IGOs did not have an appreciable effect and may have had a slightly negative effect over time.

The treatment group's increased performance in the second 70-minute interval followed by a decrease in performance during the final test interval was replicated in the cross-over group. The cross-over group was tested twice, with two months between testing sessions and IGOs included in only one session (Andrew et al., 2003). A possible explanation is that IGOs, when combined with long periods of concentration, create additional mental workload and fatigue, resulting in poorer performance. This trend has not been found with X-ray screeners who typically view X-rays for 20-minute periods before rotating to another function. The working period in excess of three hours for operators in the Andrew et al. (2003) study may have induced fatigue and lowered ultimate detection rates despite short breaks. This highlights the importance of effective shift management. Ensuring that operators take regular breaks during monitoring activities appears to be essential in managing work demands over the extended shift period of CCTV operations. This may become even more important with the added demands of IGO identification.

To examine the results further in the Andrew et al. (2003) study, the treatment group was divided into high performers who detected high numbers of significant events, average performers, and low performers, and the analysis was repeated. IGOs did not have an appreciable or significant effect on high or low performers, but increased performance at the end of the second period appreciably for medium performers although this was not significant. This suggests that personnel with different capacities may react differently to IGO interventions. Higher performers may continue to perform well, low performers may improve only slightly, and medium (the majority) performers may benefit most from the intervention.

The results of the Andrew et al. (2003) study are not statistically significant but suggest that IGOs may be useful for CCTV. While it is possible that IGOs may not be appropriate to CCTV, one explanation for the lack of significance is that IGO design requires more careful consideration. In this context, 'design' refers to an optimal combination of IGO characteristics and an effective implementation strategy for inserting IGOs into displays. The focus of the current research is on identifying IGO characteristics that enhance the detection of significant events. Although issues regarding IGO implementation are not the main focus of this research but are discussed in terms of future research directions in chapter 8.

3.2. Intervention in the Current Study

IGO systems could fulfill several different objectives and each has implications for the characteristics of IGOs. The key objectives of IGO systems are to reduce the vigilance decrement, maintain vigilance and motivation, to assess whether operators are looking at the display in order to maintain detection rates for significant events, and enhance the detection of significant events (Andrew et al., 2003). These objectives are adopted in the current research.

The first objective of enhancing vigilance and motivation is based on manipulating signal frequency. Detection performance is related to signal frequency (Baker, 1960; Balakrishnan, 1998; Koelega, 1996; Mackie, 1987). By introducing IGOs, the frequency of events to which operators are required to respond is increased. Expectancies regarding event frequency impact on observation and vigilance performance (See et al., 1995). This is indicated by arousal and expectancy theories of vigilance (Loeb & Alluisi, 1984) and SDT (MacMillan & Creelman, 1991).

Arousal theories explain performance improvements through the maintenance of appropriate arousal levels during the vigilance task (Matthews & Davies, 1998). Infrequent significant events decrease arousal below optimal levels (Koelega, 1996). The insertion of IGOs arguably alters arousal levels. Based on arousal theory, IGOs should be inserted at a frequency that optimises arousal. However, learning and neurological theories predict that responses are habituated over time (Davies & Parasuraman, 1982). Therefore the benefits of IGOs may attenuate over the shift.

SDT proposes that alterations in expectancies lead to a change in the observer's response criterion (MacMillan & Creelman, 1991). Altering the frequency of significant events influences expectancies regarding events, and these in turn change the response criterion and detection performance (Long & Waag, 1981). This suggests that the insertion of IGOs increases expectancies regarding the frequency of events requiring a response. These in turn should alter the response criterion so that operators are more likely to decide that an IGO is present.

Multiple resource attention theory predicts that tasks requiring the same resources deplete resources through demand overload (Wickens, 2008). Because the detection of IGOs and significant events use the same resources, the introduction of IGOs could lead to a decrease rather than an increase in detection levels. However, the resource demand model is most applicable 'in the high demand multi-task environment, typical of any vehicle driver, overworked secretary, or commander in an emergency operations mode' (Wickens, 2002, p. 167) rather than situations where there are very few significant events, such as many vigilance tasks.

In many industries with proactive CCTV surveillance, significant events occur infrequently (Donald, 2001). Therefore the resource demand model may be less applicable to IGOs in these situations. It is argued that with CCTV systems where significant events occur infrequently, IGOs would seldom be competing with significant events for resources. Despite the potential increase in resource demand created by IGOs, it is proposed that IGOs could assist in maintaining motivation and reducing the negative effects of low signal frequency, expectancies and a conservative response criterion.

In addition to the frequency of significant events and amount of resources demanded by the tasks, the demand on resources is also modulated by preemption, task difficulty and the resource allocation policy (Wickens, 2002, 2008). It is proposed that the characteristics and design of IGOs influence the overall demand on resources. This is discussed in chapter 4.

The second objective of IGO systems is to assess whether operators are looking at the display in order to maintain detection rates for significant events. It is based on the premise that looking at displays is a prerequisite for and the detection of significant events. This objective has limitations in enhancing the detection of significant events. Directing one's gaze at displays does not guarantee that one is purposefully searching the display or that a significant event will be recognised when it occurs (McCarley, Kramer, Wickens, Vidoni, & Boot, 2004). In addition, the detection of IGOs and significant events are not necessarily related. While this may hold true in some cases, it cannot be assumed for all cases.

The potentially beneficial effects of IGOs on expectancies may be restricted to the detection of IGOs and may not generalise to significant events. In this situation, IGO detection rates could be high yet significant events could be missed. Several factors are likely to influence the relationship between IGO and significant event detection. These include the attention sets required for IGOs and for significant events, the potential for IGOs to distract operators from significant events, different detection difficulty levels for IGOs and significant events, and different consequences for IGO and significant event detection in the operational context. These factors should also be considered in IGO design.

Studies that only measure IGO detection and not the detection of significant events are seen as being problematic (e.g., Neil et al., 2007) as they assume that IGO and significant event detection are related. Similarly, the use of IGOs in evaluating and managing performance should be based on a demonstrated relationship between the detection of IGOs and the detection of significant events. Where a relationship exists, IGOs can be used to measure and predict performance regarding the detection of significant events and provide feedback to operators (Donald & Donald, 2008).

The third objective of IGOs systems is to use IGOs not only to heighten vigilance or ensure that operators are gazing at the display, but to enhance the process of recognising and detecting significant events (Andrew et al., 2003). This objective is of most interest in the current research. It aims to use IGOs to assist operators in guiding attention to significant events on a goal-driven basis and deploying attention resources to them. It is contended that this can be achieved through appropriately designed IGOs that assist in developing and activating SA, mental schemas and attention sets that are relevant to both IGOs and significant events. Not all IGOs are likely to achieve this objective and therefore the identification of appropriate IGO characteristics is important. This objective shares similarities with the use of pictorial symbols used to enhance the processing of safety information (e.g., Hancock et al., 2004).

The focus of the current research is on how IGOs can be designed to enhance the recognition and detection of significant events and the characteristics required to achieve this. IGOs could potentially include an unlimited range of content and characteristics. However, not all of them are likely to be effective. The type of objects and events portrayed by IGOs and how they are depicted and implemented, need to be matched to the particular objectives of the IGO system (Donald & Donald, 2008).

Using IGOs to enhance the detection of significant events is a significant departure from traditional ASI and TIP studies and practice, and goes beyond manipulating signal frequency and expectancies (Donald & Donald, 2008). The emphasis on attention processes contrasts with many of the strategies that have been proposed to improve vigilance. Over two decades ago Craig (1984) noted that few interventions aimed at improving vigilance performance had been tested or attempted to address attention processes. A review of research on vigilance indicates that this comment still applies today (Donald & Donald, 2008). Interventions have typically focused on increasing motivation and altering the stress and fatigue levels of operators by providing appropriate work/rest schedules (Craig, 1984). These could be applied to the CCTV context by redesigning work so that monitoring is alternated with other tasks and monitoring is performed for shorter periods. However, these interventions alone are unlikely to address all performance issues.

Training is one of the few interventions that addresses attention processes by developing SA and an understanding of what constitutes significant events (Donald, C., 2008). This knowledge can be translated into mental models of significant events to assist operators in guiding attention to these events. The current intervention also makes use of SA, mental schemas and attention sets in guiding attention, but does not involve formal training. Rather, these processes are developed and reinforced exclusively through the use of IGOs.

The IGOs are intended to assist attention guidance to significant events. However, it is proposed that their ability to do this is related to how they are designed and related to significant events. While graphics with certain characteristics may enhance the detection of significant events, other types of graphics may distract operators from potential significant events and could be detrimental to the detection of significant events. Therefore it is important to identify characteristics that enhance the detection of significant events. First, however, it is necessary to clarify terminology regarding images and how IGOs compare with graphics in other contexts.

3.3. Types of Images and Contexts

Different types of images are used in a number of contexts in addition to TIP images, such as safety and warning signs and labels, computer user interfaces and transport systems. An understanding of the terms used and the various functions of images indicates how IGOs compare with other types of graphics and where theory on the characteristics of graphics is likely to inform IGO design. Several different terms are used to refer to graphic images, including icons, signs, indices, symbols, pictograms, and pictorials. These terms have been defined in various ways with the key difference being their relationship to the object they represent (McDougall et al., 2006).

An 'Icon' originally referred to a pictorial representation that physically resembles the object it represents (Familant & Detweiler, 1993). Used in this way, icons have a direct relationship with their referents. A 'sign' or 'index' has an indirect relationship with its referent and signifies, indicates or 'points to' the object or event it represents. A sign has some resemblance to or relationship with the object is represents, but this is not as direct as with an icon. Learning is often necessary to understand the association between the sign and its referent (McDougall & Isherwood, 2009). A symbol bears no physical relationship with its referent and the relationship between the two is usually arbitrary (Familant & Detweiler, 1993). Like signs, a symbol's meaning is learnt and requires interpretation. IGOs could potentially be icons, signs or symbols, and it is important to consider which of these is likely to be most useful in assisting the detection of significant events.

Despite the original differences between icons, signs or indices and symbols, their meaning has altered over time. Icons now include abstract symbols and need not depict real objects (Familant & Detweiler, 1993). Therefore 'The term "icon" has been generalised to mean any kind of largely pictorial and potentially meaningful expression' (Familant & Detweiler, 1993, p. 710). More recently, it has been used to refer to signs and symbols (e.g., McDougall et al., 2006). In the computer context, icons are usually small images (Thatcher et al., 2005). They are used to 'help individuals interact with machines and their environment' and include a range of images with different relationships to the objects or functions they are intended to represent (McDougall & Isherwood, 2009, p. 325). The terms 'pictogram' and 'pictorial' are also relevant to the current research. These refer to 'simple diagrams [that] are used to improve understanding of concepts' (Yin et al., 2008, p. 815). They may include icons, signs and symbols based on their original definitions.

In the current research, the terms are used in particular ways. 'Graphic' and 'image' refer generically to all types of diagrams used to convey information regardless of their relationships with their referents. Consistent with recent use of the term, 'icon' refers to small graphics used to assist human-machine interaction, including computer-user interfaces. In this study, an icon need not physically resemble its referent and is therefore used in a different way from its original meaning. Graphics with a strong relationship to their referent are referred to as 'concrete images'. 'Sign,' 'pictorial' and 'pictogram' are reserved for graphics used in warning and safety communications, such as on notice boards, labels, and brochures. 'Index' and 'symbol' are used in their original definitions, with indices having an indirect relationship with their referents, and 'symbols' having arbitrary relationships to their referents. IGOs are defined as any graphic image that is inserted electronically into a CCTV video stream. They could be icons, signs, indices, symbols, or pictorials. However, they differ from graphics in certain other situations by virtue of their function and context.

The purpose and instrumentality of graphics varies across contexts. With computer-user interfaces, icons represent objects or functions which are activated immediately by clicking on the icon. Icons are highly instrumental in activating computer functions and there is a causal link between the icon and the computer's response. Feedback regarding whether the choice of icon is correct is inherent in the process because the intended function either is or is not initiated. This reinforces correct choices of icons and facilitates learning and understanding of what the icon represents.

Safety signs have less instrumentality than icons. They are intended to convey information and change behaviour (Akerboom, 1994). To achieve these objectives, the user needs to understand the information the sign conveys (Cho et al., 2008). Comprehension is a key moderator of the safety sign's effectiveness and is related to the sign's characteristics (Davies, Haines, Norris, & Wilson, 1998).

IGOs are intended to assist the search and recognition of observable events. However, the causal link between IGOs and significant events is weak as it is not known what or when significant events will occur. With IGOs, there is no direct causal link between the graphic and significant events. This is likely to affect the learning of IGOs' meanings and their ability to assist the detection of significant events. Therefore it is important that the link between IGOs and significant events is understood easily by operators. IGO characteristics that facilitate comprehension of this link therefore need to be identified.

The usefulness of warning and safety pictorials in influencing attitudes, beliefs, motivation and behaviour, is related to the degree to which they are understood by users (Wogalter & Laughery, 1996). The pictorials' characteristics are crucial to comprehension (Dowse & Ehlers, 2004; Houts, Doak, Doak, & Loscalzo, 2006; Mansoor & Dowse, 2007; Miller & Stanney, 1997; Wilkinson, Cary, Barr, & Reynolds, 1997). Therefore IGO characteristics will influence operators' understanding of their relationship to significant events. IGO design is important in order to avoid misinterpretation. The meaning of pictorials is not governed by clear conventions that convey meaning (Isherwood et al., 2007). Users act as interpreters of images, leaving room for errors (Hancock et al., 2004). Therefore it is important to design IGOs to facilitate accurate comprehension.

As with IGOs, certain safety signs are intended to encourage visual search. For example, a road sign warning of children crossing a road reminds drivers to look out for children. Similarly, the purpose of IGOs is to encourage the search for particular events. However, safety signs and pictograms often cue behaviour in a different way from IGOs. Many safety signs are linked to specific geographical locations or events. For example, 'stop' signs are found at intersections and directly cue drivers to stop. Pictograms on medicine dosage directions act as cues regarding how medicine should be taken. IGOs, on the other hand, do not cue significant events in the same way. With CCTV there is far greater variety regarding the significant events that could occur, the forms they could adopt, and high unpredictability regarding the frequency of occurrence. These factors create a complex search task with greater uncertainty. IGOs are merely guides rather than prescriptions or cues for specific behaviours. They are intended to maintain a set of cognitive 'behaviours' (i.e., visual attention) that assists in the related behaviour of visual detection.

Given the low instrumentality of IGOs and the fact that they do not cue significant events, comprehension of their link to significant events is likely to be a key aspect in their effectiveness. Theory regarding the design of graphics in other contexts has been developed and is used to identify a range of characteristics that could be applied to IGOs. In addition, theory regarding graphic characteristics that facilitate understanding of the meaning of the graphic is useful for IGO design. However, existing literature contributes relatively little to the key issue of how IGOs can be designed to facilitate the search for and recognition of significant events.

The current research therefore has several aims. Firstly, it aims to contribute to the development of an intervention to enhance the proactive detection of significant events by inserting graphic objects into live CCTV footage. The second aim is to identify a range of characteristics to consider in IGO design that influence attention and search processes and the detection of significant events. This is achieved by drawing from visual attention theory and icon characteristic theory. IGO design includes characteristics that influence the content of IGOs and how this is depicted, as well as how IGOs are inserted into displays, or an implementation strategy. Underlying processes whereby IGOs could assist the detection of significant events are discussed. The third and more specific aim is to evaluate the contribution of one of the characteristics, semantic distance, to the detection of significant events. The latter is examined empirically.

The next chapter discusses processes whereby IGOs are seen as influencing attention processes in the detection of significant events. The chapter thereafter identifies a range of IGO characteristics and selects certain of these for the empirical study.

CHAPTER 4: IGOs AND THE DETECTION PROCESS

As with artificial signals and TIP, IGOs are envisaged as reducing monotony and increasing operator motivation, arousal and alertness. This is based on arousal and expectancy theories and the finding that the frequency of significant events is related to the vigilance decrement (Van Wert et al., 2009). While this may be beneficial, it is unlikely to be sufficient in terms of facilitating the detection of significant events rather than merely detecting the IGOs. Therefore it is necessary to examine processes whereby IGOs could facilitate the detection of significant events, rather than merely drawing attention to themselves.

It is proposed that IGOs affect various attention processes involved in visual search and the recognition of significant events. Mechanisms proposed for this include: the creation of specific attention sets; biasing the attention filter to assist selective attention; assisting top-down and bottom-up attention guidance; and influencing the schemas contained in working memory. Discussion of these mechanisms is intended to provide an understanding of the underlying processes whereby IGOs could assist the detection of significant events. Specific IGO characteristics which are most likely to achieve this are discussed in the next chapter. However, it is noted at this point that IGOs with different designs and characteristics are likely to vary in their ability to influence the detection of significant events in a positive manner. The following discussion is based on the assumption that IGOs are designed in an optimal manner.

Attention processes form the core of visual search, scanning, information sampling, perception, and the recognition of significant events (Donald, 2001). The term 'attention' typically refers to a capacity which can be applied voluntarily to a situation (Pashler, 2008). Some authors refer to attention as a 'fuel' or a resource (e.g., Wickens & McCarley, 2008). However, researchers have used the term to refer to various different processes, including aspects of human cognition that individuals can control, those related to a limited capacity, and ways of dealing with these constraints (Shiffrin, 1988). Attention is recognised as consisting of multiple processes rather than a single entity (Parasuraman, 1998). Agreement on a single definition of attention is unlikely to be reached. An understanding of the processes involved is likely to be more useful. Attention processes are incorporated into the discussion below on how IGOs could influence the detection of significant events. Attention processes were discussed in chapter 2 in relation to vigilance processes and are incorporated into the discussion below that focuses more specifically on how IGOs could influence the detection of significant events.

With CCTV, the requirement that operators monitor multiple complex displays (and areas within scenes) raises the question as to the types of attention that operators use and how they allocate shared attention resources across displays. The types of attention are focused (similar to concentration), selective (choosing certain stimuli from others), divided (sharing attention between concurrent tasks), switched (changing attention from one task to another and back again), and sustained (maintaining attention over an extended period of time) (Wickens & McCarley, 2008). All types of attention are relevant to operators, who are required to select relevant significant events, focus attention on events, divide or share attention between objects, people and areas, switch attention between cameras, objects and displays, and sustain attention over lengthy periods. How operators deploy these types of attention will affect the viewing process. For instance, they may rapidly switch attention between different objects, focus on one object, or try to include several objects in their gaze simultaneously. This has implications for what they do and do not perceive and therefore influences performance.

Attention processes and eye movements are closely related (Wickens & McCarley, 2008). Rapid eye movements (saccades) generally indicate the deployment of attention and information sampling (Rolfs, Jonikaitis, Deubel, & Cavanagh, 2010). Although this is not always true (covert attention may be deployed), the correlation between eye movements and attention is large enough to be accepted as valid in work-related contexts (Wickens & McCarley, 2008). Thus 'what we see, understand and remember from the visual world is tightly tied to where our eyes are pointed' (Henderson, 2007, p. 22). This is particularly relevant to surveillance operators whose key function is to perform visual search and detection. Ways in which IGOs could influence eye movements and attention are therefore important in the current research. The first mechanism whereby this is proposed to occur is that of the attention set.

4.1. Mental Models, Schemata and Situation Awareness (SA)

The role of expectancies in vigilance performance was discussed in chapter 2. In addition to vigilance tasks, expectancies play an important role in visual search (MacMillan & Creelman, 1991). They have been included in various models and theories of visual search, such as the SEEV model (Wickens & McCarley, 2008), the revised feature integration theory (Treisman & Sato, 1990), and guided search theory (Wolfe, 1998). Expectancies are incorporated into cognitive schemata, mental models and attention sets, all of which influence search, information processing and the recognition of significant events. IGOs are envisaged as influencing these mechanisms and if designed appropriately, could be used to develop these mechanisms to enhance detection.

The use of IGOs in this way is based on the use of graphic images in other contexts such as safety signs and graphical user interfaces. Appropriately designed images influence the types of search patterns and mental models that users develop (McDougall, Curry, & de Bruijn, 2001), how effectively information is communicated, and can reduce misinterpretation (e.g., Isherwood et al., 2007; Laughery, 2006; Mansoor & Dowse, 2007; McDougall & Isherwood, 2009; Miller & Stanney, 1997; Pappachan & Ziefle, 2008). This suggests that in the CCTV context, it is important to consider how IGOs could best be designed to assist the search process through the development of mental models and attention sets for significant events.

Both schemata and mental models are cognitive representations of concepts, objects or events (Miller & Stanney, 1997). These terms are used interchangeably by some authors (e.g., Endsley, 2000) and in the current research. Schemata represent underlying knowledge about concepts, sequences of actions and events. These representations are fairly permanent, stored in long term memory and activated when required (Miller & Stanney, 1997).

In the context of information sampling, a mental model is ‘a set of expectancies about how frequently and when events will occur on each channel and about the correlation between events on pairs of channels' (Wickens, 1992, p. 78). Thus mental models include beliefs about the occurrence of events and factors that are related to them, including pre-conceived ideas, stereotypes and knowledge. Mental models can also refer to an understanding of how a device or system works (e.g., Kieras & Bovair, 1984). This is reflected in Carroll and Olson's (1987, p. 12) definition:

> … a rich and elaborate structure, reflecting the user's understanding of what the system contains, how it works, and why it works that way. It can be conceived as knowledge about the system sufficient to permit the user to mentally try out actions before choosing one to execute.

This 'mental picture' or 'description' reflects people's understanding of events and systems, informs expectancies regarding when, where and how events occur and influences visual sampling and scanning (Rauschberger & Yantis, 2006). Mental models assist comprehension of what is occurring and allow options for responses to be considered and future events predicted (Johnson-Laird, 1980). In the current research, mental models include aspects such as knowledge and understanding of the CCTV system, (e.g., camera positions relative to geographical layout) and significant events (e.g., how significant events are set up and executed, where and when they occur, expectancies and profiling of perpetrators and victims), and responses required to different events.

To detect and recognise visual stimuli, relevant schemata need to be activated (Familant & Detweiler, 1993). Schemata and mental models assist in organising knowledge and perceptions and enable information to be processed quickly (Johnson-Laird, 1983). Visual stimuli that trigger schemata are processed more quickly than those for which no relevant schema exists (Theeuwes, 1996). Schemata assist in dealing with information overload at the bottleneck in information processing that is generally accepted to occur (Rauschberger & Yantis, 2006).

Schemata and mental models form an integral part of SA. SA has been defined in various ways (Kirlik & Strauss 2006, O'Brien & O'Hare, 2007; Sinclair, 2007), but Endsley's (1988) qualitative model is adopted for this study. SA is defined as 'the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future' (Endsley, 2000, p. 3). The concept of SA is controversial. Its detractors criticise it as a folk model which does not meet the standards required for scientific investigation (Dekker & Hollnagel, 2004). While critics acknowledge that it is useful in describing phenomena, they caution that it should not be used as a causal explanation (Flach, 1995). Despite these issues, research on SA has proliferated (Endsley, 2000). In the current research where there is little prior knowledge of how operators view displays, SA has the potential to assist in describing the viewing and detection process and to contribute to an understanding of the cognitive processes involved.

SA is related to what observers perceive, the accuracy of the mental picture they form of situations, how they comprehend and interpret situations, and the extent to which they are able to use the data to predict situations and make appropriate decisions (Endsley, 2000). Accurate SA is critical in making correct decisions (Wright, Taekman, & Endsley, 2004). Operators with high levels of SA are more likely to perceive, recognise, and comprehend significant events than those with low levels of SA (Federico, 1995). The current research proposes that IGOs can be used in the development of SA and influences the search process.

SA develops from a number of direct, indirect and geographically distributed information sources (Endsley, 2000). These include salient stimuli (e.g., CCTV displays) and cognitive knowledge structures such as goals, expectations, mental models, learnt scan patterns, information sampling strategies, and previously acquired data. Knowledge from these sources is embodied in mental models that guide attention to relevant events and states (Nunes & Mogford, 2003).

In the current research, it is proposed that appropriately designed IGOs be one of the information sources used to develop and reinforce SA relating to significant events. By graphically depicting aspects of significant events, IGOs could contribute to the development and reinforcement of accurate schemata and mental models of significant events. Accurate mental models assist in the performance of tasks (Miller & Stanney, 1997).

Appropriately designed IGOs could reflect pertinent aspects that operators should watch for. This is similar to the use of graphic images in icons, signs and symbols in other contexts. Graphics have been used in computer-user interfaces and safety signs in road and transport systems (e.g., Laughery, 2006; McDougall et al., 2001). In these contexts, images activate relevant schemata and mental models, allowing for comprehension and use of the information in the image (Miller & Stanney, 1997). A similar process could apply to IGOs. The ability of IGOs to assist in developing mental models of significant events is seen as being dependent on IGO design. This is based on the finding that the nature of image content influences the mental models that are developed and used in graphical user interfaces (McDougall et al., 2001; Miller & Stanney, 1997).

4.2. Attention Sets and Selective Attention

SA has a major impact on search and detection, influencing what is seen and noticed, whether events are recognised, and the type of responses made (Endsley, 2000). A key part of this is the attention set which is defined as 'A preparatory state of the information processing system that prioritizes stimuli for selection' (Leber & Egeth, 2006, p. 565). It consists of goals, expectancies and values (Wickens & McCarley, 2008) and creates 'a perceptual readiness to locate [a target]' (Most, Scholl, Clifford, & Simons, 2005, p. 218). Simply stated, the attention set reflects what the observer is looking for and has a major impact on what is perceived and noticed and what is not selected for attention. The attention set is similar to SA in that both processes provide conditions to guide how people search their environments. However, the term 'attention set' refers more specifically to the bias towards specific information (Pashler, 1999), or particular events that people are searching for at a specific time. SA is a broader concept that refers to a number of processes (e.g., mental models), one of which is the attention set.

Stimuli that are similar to the observer's attention set are more likely to draw attention than those that are dissimilar (e.g., Most, Simons, Scholl, & Jimenez, 2001). Change and inattentional blindness studies lend further support to the importance of the attention set as observers frequently fail to notice events they are not looking for (Cole & Liversedge, 2006; Gibson & Peterson, 2001). This suggests that IGOs could play a major role in shaping operators' attention sets, what they look for and what they detect. This is particularly relevant to significant events that occur infrequently, unpredictability and which have low conspicuity (Donald & Donald, 2008).

Visual stimuli that are included in the attention set are more likely to be noticed and recognised, while those that are not included may not be noticed, even if they are above the threshold of visibility and within focal vision (Most et al., 2001). Operators who adopt attention sets for significant events are more likely to perceive and detect significant events than those who do not. Where IGOs are used, operators require attention sets for both IGOs and significant events. If IGOs are designed in a way that requires a different attention set from significant events, mental workload and attention resources required for detection could be increased. This could also result in high detection rates for IGOs and low rates for significant events. Therefore it seems intuitively logical to design IGOs so that they share a common attention set with significant events. This is assumed to limit the increase in mental workload and attention resources. The focus of the current research is therefore on identifying IGO characteristics that allow IGOs to maintain, consolidate and enhance attention sets for significant events.

It is envisaged that appropriately designed IGOs will assist in creating and maintaining schemata, mental models, and attention sets for significant events. By inserting IGOs, mental models can be developed and maintained in working memory. These would create 'readiness' to perceive and recognise similar behaviours and events on CCTV displays. This rests on the assumption that IGOs contain certain characteristics and that not all IGOs would assist in the detection of significant events. Characteristics most likely to assist this process are discussed in the next chapter.

The attention set is one of the mechanisms involved in assisting people in selecting stimuli for the allocation of attention and overcoming a limited information processing capacity (Leber & Egeth, 2006). At any one time, too many stimuli in the environment impinge on the senses to be processed simultaneously (Pashler, 1999). Therefore selective attention chooses certain stimuli for attention and rejects others (Monsell & Driver, 2000). A filter has been proposed as the mechanism for selecting stimuli (Broadbent, 1958). The current research proposes the use of IGOs in biasing the filter towards significant events, thereby assisting selective attention and detection. It is envisaged that this could be achieved through appropriately designed IGOs that positively influence the attention set, mental models and SA.

The filter is biased to select certain information and reject others (Loeb & Alluisi, 1984). The way in which it is biased depends on perceptual characteristics present in the visual environment and the observer's attention set (Most et al., 2005; Wickens & McCarley, 2008). IGOs should be designed to maximise their potential to bias the filter towards significant events and not merely towards themselves. IGOs should therefore reflect perceptual characteristics most likely to be selected by the filter as well as content that enhances attention sets for significant events. It is envisaged that this will assist in creating attention sets that can be applied to significant events as well as IGOs. Without this commonality, different attention sets may be required for IGOs and significant events, creating extra demand on attention resources.

It is emphasised that IGOs' ability to assist detection through selection attention, mental models, and attention sets is seen as being dependent on IGO design and characteristics. This is the central focus of the current research and specific characteristics are discussed in chapter 5.

4.3. Top-down and Bottom-up Attention Guidance

In addition to assisting in the development and maintenance of SA and attention sets, it is proposed that IGOs could assist attention guidance through displays in order to find significant events. However, it is noted that IGOs are not intended to function as cues that guide attention on a spatial basis. This is because the location of significant events is uncertain although contextual cues may suggest where they are likely to occur. Rather, IGOs are seen as guiding attention by providing graphic depictions of relevant visual features of significant events and by activating relevant mental models and attention sets. These processes correspond to the two key processes involved in attention control.

The first attention guidance process is based on physical properties inherent in visual objects which tend to draw attention, even if they are not related to the observer's goals (Leblanc & Jolicoeur, 2005). It is usually seen as being involuntary and is referred to as stimulus-driven, bottom-up or exogenous attention control (Folk & Remington, 1996). The second process is guided by factors under the observer's control, including goals and attention sets (Chen & Zelinsky, 2006). This is known as goal-directed, top-down or endogenous control. An understanding of these processes indicates how IGOs could assist attention guidance. As with all instances of perception (Serences, Shomstein, Leber, Golay, Egeth, & Yantis, 2005), both processes operate in the CCTV context. Search is influenced by the physical characteristics of stimuli or displays and the goal of detecting significant events. Where IGOs are used, it can also be assumed that the physical characteristics of IGOs influence the search and that operators have the additional goal of detecting IGOs.

Visual properties that draw attention in scenes on a stimulus-driven basis are visual salience (Pomplun, Reingold, & Shen, 2003), new objects (Brockmole & Henderson, 2005a; Fleetwood, & Byrne, 2006; Itti & Koch, 2000; Irwin Colcombe, Kramer, & Hahn, 2000), proximity (Fleetwood & Byrne, 2006), item interrelationships (van Zoest, Giesbrecht, Enns, & Kingstone, 2006), and certain types of motion (Chastain, Cheal, & Kuskova, 2002; Franconeri, & Simons, 2003). Stimuli that differ from their surrounds and 'pop out' (high visual salience, abrupt onsets, luminance change, new objects) tend to be detected efficiently relative to those that receive attention on a more goal-directed basis (Bauer, Jolicoeur, & Cowan, 1996a; Itti & Koch, 2000). IGOs with these physical properties are likely to be detected quickly relative to those without these properties. While this may assist in the detection of the IGOs themselves, it may not be sufficient in using IGOs to assist in the detection of significant events. Goal-directed attention provides more indicators as to how this could be achieved.

Goal-directed attention refers to 'the observer's ability to control what regions or objects in the visual field are selected for further visual processing given a set of goals or beliefs about the current task' (Yantis, 1993, p. 676). Goal-directed attention is voluntary, intentional (Serences et al., 2005), and is usually associated with searching for objects and performing tasks. Therefore it assists operators in their detection task.

A broad range of cognitive knowledge structures guide search through scenes in a goal-directed manner. These include expectancies (Evans & Treisman, 2005), task goals and demands (Einhaüser, Rutishauser, & Koch, 2008; Yantis, 1996), attention sets (Most et al., 2005), SA and mental models (Eimer, Nattkemper, Schröger, & Prinz, 1996), knowledge and expectancies regarding areas providing maximum task-relevant information (Castelhano &

Heaven, 2010; Gilchrist & Harvey, 2000; Pomplun et al., 2003), learnt contextual cues, semantic interpretation (Gordon, 2004; Henderson et al., 2007; Najemnik & Geisler, 2005; Torralba, Oliva, Castelhano, & Henderson, 2006), habit (Foulsham & Underwood, 2008), and memory (Henderson et al., 2007).

IGOs with appropriate characteristics are seen as assisting in the development and maintenance of goal-directed attention guidance. This is important because there is increasing evidence that top-down processing has a stronger influence on where observers direct their attention than visual saliency in real world scenes (Bacon & Egeth, 1994; Brockmole & Henderson, 2005b; Einhäuser, Rutishauser, & Koch, 2008; Henderson, 2003; Foulsham & Underwood, 2008; Henderson et al., 2007; Li, van Rullen, Koch, & Perona, 2005).

The dominant role of goal-directed processing is attributed to the meaningfulness of real world objects and weightings assigned to goal-directed processes (Chen & Zelinsky, 2006). The eyes tend to fixate on areas considered informative or which contain unexpected objects, enabling further processing and understanding of the scene (Brockmole & Henderson, 2005b). This is consistent with the importance of meaning and the purposive nature of searching in the real world (Henderson, 2007).

It seems logical to design IGOs so that they facilitate goal-directed processing, tie in with the goals of detecting significant events, and assist operators in deciphering visual stimuli that signify potential significant events. It is proposed that this be achieved by designing IGOs that reflect pertinent aspects of significant events, thereby developing and maintaining mental models and attention sets that are relevant to significant events, as discussed previously. Despite the likelihood that goal-directed guidance is dominant over stimulus-driven guidance in CCTV, these processes interact and are both relevant to CCTV. This is supported by the finding that gaze and attention are controlled by the interaction of bottom-up and top-down attention guidance (Duncan, 1984; Treisman & Sato, 1990; Vecera & Farah, 1994).

The physical characteristics of significant events and IGOs influence the extent to which they draw attention and how easily they are detected. For instance, a movement that is different from the pattern of surrounding movements, or the entry of a person into an area where people are prohibited could draw attention on a largely stimulus-driven basis and indicate that a significant event is occurring. Similarly, the operator's goals of detecting IGOs and significant events influence what s/he looks for, areas considered to be of interest, and decisions regarding who or what to monitor (Norris & Armstrong, 1999) on a goal-directed basis.

A number of factors modulate the relationship between bottom-up and top-down control. These include target- non-target similarity (Proulx & Egeth, 2006), the availability of attention resources (Boot, Brockmole, & Simons, 2005), the type of task involved (Bravo & Nakayama, 1992), long-term memory representations of similar tasks, spatial and temporal certainty (Wickens & McCarley, 2008), perceptual organisation and grouping, working memory representations (including task goals and features defining the target), and expectations about non-targets (Yantis, 2000). The effectiveness of IGOs in facilitating the detection of significant events may be influenced by these factors.

The dominance of goal-directed and stimulus-driven processes is also related to knowledge about targets. Top-down attention control can be used when target characteristics are known. However, when target knowledge is unavailable, stimulus properties are likely to be given more weight (Wickens & McCarley, 2008). This is similar to the finding that stimulus-driven guidance is likely to be dominant when goal-directed processes are unavailable (Chen & Zelinsky, 2006). Although operators performing proactive surveillance may not know specific details of significant events that may take place, they have mental models and expectancies regarding the types of significant events that could occur. These inform their search. Mental models are stored in long term memory which therefore plays a role in search and detection.

4.4. Long Term and Working Memory

It is envisaged that visual stimuli provided by IGOs assist in activating and maintaining relevant attention sets and mental models of significant events in working memory. This is similar to the use of images in safety warnings which remind users of safety issues and bring latent knowledge from long term memory into conscious awareness (Laughery, 2006).

Mental models function as templates for events and need to be activated for recognition to occur (Familant & Detweiler, 1993). This takes place in working memory which is defined as 'A system responsible for the active maintenance of goal-relevant information in the face of concurrent processing and/or interference' (Conway & Kane, 2001, p. 365). Thus it allows for both storing and processing of material during complex tasks (Wickens & McCarley, 2008).

Working memory assists selective attention by facilitating the selection of stimuli that match object representations, thereby enhancing top-down control of attention and the achievement of goals (Soto, Heinke, Humphreys, & Blanco, 2005). Templates in working memory are associated with both costs and gains in search efficiency. When the template is valid, search is faster, especially at larger display sizes. Thus working memory primes the target (Wickens & McCarley, 2008). However, when the template is not valid, search is slower and a cost is involved. This suggests that IGOs should assist in the creation of valid templates for significant events if they are to assist in the detection of those events. Such IGOs would reflect key features of significant events rather than be unrelated to them. This is discussed further in the chapter 5.

Theory suggests that IGOs need not be exact replicas of significant events in order to activate relevant mental models of the latter. This is based on the finding that templates in working memory need not be replicas of targets, but should merely be associated with targets for priming to occur (Moores, Laiti, & Chelazzi, 2003). The material used in IGOs should trigger templates that are pertinent to significant events, in order to prime operators for significant events. Therefore designers should give careful consideration to the material represented in IGOs and its relationship to significant events.

When IGOs are not related to significant events, different mental models are likely to develop for IGOs and significant events. In this situation, two disparate sets of mental models and attention sets would need to be activated and maintained in working memory. Based on multiple resource attention theory (Wickens, 2002), these are likely to compete for and deplete attention resources.

4.5. Potential Disadvantages of IGOs

While it is envisaged that appropriately designed IGOs will assist the detection of significant events, it is possible that IGOs could have a negative effect on performance. This could be due to fatigue, the depletion of attention resources, increased mental workload, and distraction. All these factors have been related to decreased vigilance performance (e.g., Loeb & Alluisi, 1984; Sawin & Scerbo, 1995; Szalma, Warm, Matthews, Dember, Weiler, Meier, & Eggemeier, 2004; Wickens, 2002). The detection of IGOs and significant events require the same visual attention resources, and therefore the use of IGOs could deplete attention resources in a shorter time than where IGOs are not used. Similarly, mental workload and working memory could be overloaded by the use of IGOs. However, IGOs which share attention sets and mental models with significant events may be less likely to overload these systems than those that require different attention sets and mental models.

Distraction could occur by directing the gaze away from significant events or creating attention sets that compete with those required for significant events. Given the relationship between focal vision and visual attention, attention is less likely to be directed to stimuli that are not within the useful field of vision (Wickens & McCarley, 2008). When operators look at IGOs, they are therefore less likely to allocate attention resources to significant events. However, distraction by IGOs is unlikely given the relative rarity of significant events. Competition between different attention sets and mental models for IGOs and significant events has already been touched upon and will also be elaborated upon in chapter 5.

Attention theory provides a basis for explaining how IGOs could assist in the detection of significant events. However, the processes discussed rest on the assumption that IGOs are designed in a way that facilitates the direction of attention to significant events and minimises potential negative effects. The following chapter identifies characteristics that could meet these requirements.

CHAPTER 5: IGO CHARACTERISTICS

Research regarding vigilance, visual search, attention capture, and image effectiveness, provides converging evidence that the characteristics of visual stimuli influence search speed and accuracy. However, existing research contributes little to knowledge regarding characteristics of graphic images that assist the search for other events. While speed and accuracy are important in CCTV, the emphasis of the current study is on designing IGOs so that they assist the detection of significant events. IGO design in the CCTV context was considered in an initial study by Andrew et al. (2003), but specific characteristics were not explicitly identified or manipulated.

The key aim of this study is to identify IGO characteristics that enhance the recognition and detection of significant objects, events and behaviours during proactive live surveillance. While some IGOs may assist in checking whether operators are directing their gaze at the display, these IGOs are not the focus of the current study unless they are also aimed at facilitating the detection of significant events. Although the specific nature of significant events is not necessarily known in advance, certain objects, behaviours, patterns and events are often associated with significant events.

This chapter identifies a range of possible characteristics that could be used in designing IGOs and discusses how they could contribute to the detection of significant events. It then discusses how the characteristics could be combined in IGOs and the implications for the detection process. As the complete list of characteristics extends beyond the scope of this research, certain characteristics are carefully selected for the current research.

The range of characteristics that could be used in IGO design is drawn from research literature regarding images in various contexts as well as visual search and attention capture. Research from these areas provides converging support for the importance of the characteristics discussed below. These are semantic distance, concreteness, complexity, salience, novelty, dynamic properties and realism. Although research suggests that these are separate characteristics, interactions occur between them and they cannot be considered entirely in isolation (McDougall, Curry, & de Bruijn, 1999). Each characteristic is discussed separately as far as possible below. This section on characteristics concludes with a discussion of how the characteristics can be categorised conceptually and combined physically in IGOs. All characteristics exist on continua rather than being 'all or nothing' categories.

5.1. Semantic Distance

Semantic distance has major implications for IGO design as it affects comprehension, attention processes whereby IGOs are proposed to influence the detection of significant events, and the functions they represent. It is defined as 'the closeness of relationship between the icon [or image] and the function it represents' (Isherwood et al., 2007, p. 467). This is similar to the concepts of 'pertinence' (Bundesen, 1990) and 'articulatory distance' (Hutchins, Hollan, & Norman, 1985). Pertinence refers to the importance of visual elements based on their relevance to the task and identity of the target (Bundersen, 1990). Articulatory distance is the ease of inferring the meaning or function of an icon (Hutchins et al., 1985).

It is emphasised that semantic and articulatory distance and pertinence all refer to the function of the image, rather than its visual appearance. (The latter is included in concreteness which is discussed in section 5.2). This is important because the images are used to refer to something else and have no purpose without their function. The relationship between image and referent needs to be understood for the image to be effective. The relationship between image and referent function is maintained in the current research, where semantic distance is defined as the closeness of relationship between the IGO and the function of the significant event, object, or behaviour it represents.

The function of icons is often prescribed in computer user interfaces, leaving designers with the task of developing appropriate icons (McDougall & Isherwood, 2009). With CCTV, the function of each IGO is less obvious and needs to be identified by the designer. Although each IGO should have a specific function in assisting the detection of significant events, the range of material that could be used is extremely broad and should be based on a careful analysis of significant events and identification of specific features that indicate their occurrence.

Theory on semantic distance assists in narrowing the range of material that is relevant to significant events by defining different types of relationships between images and the objects or events they represent. Three types of relationships are described - direct, indirect or arbitrary (Isherwood et al., 2007). Direct relationships occur when functions are shared between the image and referent and the relationship is close (Familant & Detweiler, 1993). With indirect relationships, shared characteristics and functions are inferred, while arbitrary relationships can only be understood if users have learnt their meaning (Isherwood et al., 2007). An icon resembling a printer and the print function on a computer-user interface has a direct relationship with its referent. An image resembling a tortoise used to warn people to go slowly has an indirect and implied relationship with its referent. The yield sign (triangle) used in road signs has an arbitrary relationship with its meaning and has to be learnt to be understood.

Direct, indirect and arbitrary relationships are always based on an underlying relationship between the image's function and its referent. An additional possibility with CCTV is that IGOs could depict objects that have no referents and are entirely unrelated to significant events. Examples of these include humourous images, shapes and numbers with no significance at all. These differ from arbitrary images in that there is no underlying relationship or significance to be learnt. Semantic distance does not apply to these images as they are not signs or symbols and have no referents or meaning. These are referred to as random images in the current research. Such 'random' images were used in the Neil et al. (2007) CCTV study and a mixture of random and other images were used in the Andrew et al. (2003) study. Random images may be useful in establishing whether operators are looking at displays, but are unlikely to assist in the detection of significant events.

Semantic distance varies along a continuum, with small distances representing direct relationships and larger distances representing arbitrary relationships (McDougall et al., 1999). An examination of images used in other contexts indicates the type of relationships likely to be appropriate in CCTV. ASI and TIP use objects and images with a direct relationship to the detection task or a small semantic distance. In addition, the objects and images used have a highly realistic appearance and cannot be distinguished visually from real targets.

Given their task relevance and realistic appearance, it can be assumed that TIP images and ASI objects share the same detection difficulty levels as the objects they represent. The attention resources, scan patterns, SA (including mental models), and attention sets that assist with TIP detection also assist with the detection of actual threats for screeners. Therefore a close relationship between the detection of TIP images and real threats would be expected. Empirical research supports this as there is a relationship between TIP and real threat detection (Schwaninger, Bolfing, Halbherr, Holman, Belvayin, & Hay, 2008). This suggests that IGO content should reflect small semantic distances in order to achieve a closer relationship between IGO detection and the detection of significant events.

Depending on the material used in IGOs, the images could vary in the extent to which they represent significant behaviours or events. At the one extreme, IGOs could be random. At the other extreme, they could depict exact replicas of the threats and behaviours that operators are required to detect, as with the use of video footage of real significant events. Between these extremes are images which represent selected aspects of significant events. Given the vast array of potential IGOs, it is necessary to identify the type of relationships between IGOs and their referents that are most likely to enhance the detection of significant events. Therefore discussion of the relative dis/advantages of different semantic distances follows.

Semantic distance is related to how easily images are interpreted and understood and assists image identification (Isherwood et al., 2007). Images with small semantic distances require less learning and interpretation and their meaning is more self evident than those with large distances (McDougall & Isherwood, 2009). With CCTV, understanding what the IGO represents and the IGO's relationship to significant events is important if IGOs are to facilitate the detection of significant events. The relationship between semantic distance and comprehension suggests that IGOs with small semantic distances are most likely to be understood, and those with arbitrary relationships are least likely to be understood.

Random IGOs would not require comprehension as they are not related to significant events. This suggests that IGOs with a direct relationship to significant events should be used where possible.

In the previous chapter, IGOs were seen as assisting the detection of significant events through mental models and attention sets. This is based on the finding that the recognition and detection of visual stimuli requires the activation of relevant mental models (Familant & Detweiler, 1993). However, the ability of IGOs to activate mental models and attention sets that are relevant to significant events is argued to be related to semantic distance. IGOs with small semantic distances are likely to share mental models and attention sets with significant events due to their visual similarity and shared meaning. When IGOs with small semantic distances activate mental models, they simultaneously activate shared mental models for significant events. This should assist in guiding attention on a goal-directed basis to significant events.

IGOs with large semantic distances and random IGOs may require different mental models and attention sets from significant events. In this case, the search for IGOs and significant events could be construed as dual tasks that are performed simultaneously. Dual tasks tend to require more effort, attention resources and working memory capacity, and increase fatigue (Wickens & McCarley, 2008). This could have a negative effect on the detection of significant events. Once again, this suggests that small semantic distances are most appropriate for IGOs.

In addition to the extra demand on visual attention resources, the use of random IGOs may have several other disadvantages. Operators may prioritise the search for IGOs and neglect the search for significant events. This is supported by the singleton detection mode (Duncan, 1984) whereby observers develop an attention set for particular types of stimuli and are able to 'ignore' stimuli that do not match this attention set. It is most likely to occur with random IGOs with no relationship to significant events. Conversely, when IGOs with small semantic distances are used, similar mental models and attention sets could be applied to IGOs and significant events, reducing competition for attention resources. This could assist the detection of significant events rather than detracting from it.

Multiple resource attention theory indicates that searches for IGOs and significant events may compete for attention resources, resulting in fewer resources being available for each type of search. This is most likely to occur with arbitrary and random IGOs where the searches require different mental models and attention sets. It could decrease the detection of significant events and IGOs. Therefore random IGOs with no relationship to significant events are most likely to distract operators from the search for significant events. Multiple resource attention theory suggests that IGOs with small semantic distances are most suitable for IGOs.

The research reviewed above suggests that small semantic distances are most appropriate for IGOs. This is consistent with the types of images used in TIP. The Andrew et al. (2003) study on IGOs lends support to the recommendation that IGOs reflect small semantic distances. These researchers found that images whose content is related to significant events (e.g., a diamond) correlated more highly with the detection of significant events than those that were unrelated or random. However, their study was not designed to examine the effects of semantic distance specifically.

Consideration needs to be given to the type of material that would yield meaningful small semantic distances. This would depend on the industry and environment covered by CCTV. At an intuitive level, it seems that IGOs depicting the events or behaviours that signify the occurrence of a significant event are most pertinent to the detection task. For example, in diamond processing plants a key aim of surveillance is to prevent diamond theft. However, IGOs depicting diamonds may not be useful if the CCTV system is such that diamonds are too small to be seen on camera. Similarly, images of guns or masked people are unlikely to be helpful as these are not typically encountered in processing plant theft. Such images do not have relevant meaning or function in this context. Rather, IGOs representing behaviours indicating that diamond theft may be occurring should be used in this example as these are likely to be observed by operators.

Suitable IGOs for this example would therefore depict a person in the process of picking up something that is too small to be seen, putting something small in their mouths, flicking small objects with equipment or kicking them in order to move them to locations where they can be picked up. These behaviours are typically associated with diamond theft in processing plants. By looking for them, operators increase their chances of detecting diamond theft (Donald, C., personal communication).

The above example illustrates the point that in addition to reflecting small semantic distances, IGOs should also depict objects, behaviours or events that are observable on camera. Therefore the definition of semantic distance as applied to IGOs is extended to 'the closeness of relationship between the IGO and significant object, behaviour or event it represents, which is observable on camera and which indicates that a significant event may be occurring.' Detailed analysis of significant events is required to identify relevant, observable aspects that can be incorporated into IGOs. Significant events may take on various forms that need to be depicted. For example, stealing a diamond from a diamond processing plant often involves a number of behaviours, such as kicking or flicking the diamond to another location with a broom while sweeping, swivelling the ankle to grind the diamond into the underside of a boot, picking it up with a hand, swallowing it, or hiding it in one's uniform, amongst other actions. These behaviours can be performed in various ways. For instance, picking up a diamond can take on a number of forms that are associated with different postures. The person may be standing taking the diamond off a machine at shoulder height, or bending with his/her back to the camera and his/her hand may be visible between his/her ankles as s/he picks up the diamond.

Equipment and other items may be used to camouflage the act of picking up a diamond. A diamond can be picked up simultaneously with a hammer. Alternatively, the person could crouch down while sweeping under a machine and pick up the diamond with his/her free hand or the one holding the broom. These postures signify different forms of the same behaviour. By including them in IGOs, they could widen the operator's understanding of how diamonds could be picked up and increase the range of actions they are looking for. IGOs that are based on an analysis of how significant events occur and the behaviours and objects associated with them depict significant events accurately and have small semantic distances.

Semantic distance has been established as an important image characteristic in various contexts, but its importance changes as users become more familiar with images

(McDougall & Isherwood, 2009). Two forms of familiarity require consideration - familiarity with the pictorial depicted in the image and familiarity with its function or intended meaning (Isherwood et al., 2007). The former is related to frequency of use or experience, while the latter refers to comprehension or the relationship between the image and its referent. When familiarity with icons is low, small semantic distances are associated with faster search and greater accuracy compared with concreteness, complexity and function familiarity. However, as participants became more familiar with the images, familiarity with their function becomes more important.

Some researchers have treated familiarity as an image characteristic (e.g., McDougall & Isherwood, 2009). Although some images may be more familiar to a greater number of users than others, familiarity resides in the user rather than the image and is therefore not included as an IGO characteristic in this study.

5.2. Concreteness

Concreteness is important because it influences how referents (significant objects, behaviours or events) are represented in IGOs. It is defined as the degree to which 'real objects, material or people' are depicted (Isherwood et al., 2007, p. 466). Abstract images rely less on pictorial content and more on symbols such as arrows, shapes and lines than concrete images (McDougall et al., 1999). While semantic distance refers to the closeness between the image and its function, concreteness refers to the closeness between the image and the physical appearance of the object represented. IGOs could potentially be concrete, abstract or combine concrete and abstract features.

Concreteness contributes significantly to the understanding of an image's content (McDougall et al., 2001) and the information it is intended to convey (McDougall & Isherwood, 2009). The meaning of concrete images tends to be more self-evident than abstract images which rely more on interpretation (Familant & Detweiler, 1993) and learning (McDougall et al., 2001). Concrete images have more direct connections to objects in the real world and share characteristics with their referent objects (Familant & Detweiler, 1993). This allows people to apply their knowledge of the world to concrete images through mental representations or schemas. When people are unfamiliar with abstract images, they cannot draw on cognitive representations of objects in the real world to assist interpretation (McDougall et al., 2001). Training and familiarity, as well as text labels, can overcome the disadvantages of abstract images (Wogalter, Sojuorner, & Brelsford, 1997). Concreteness has the advantage of facilitating learning (Miller & Stanney, 1997).

Concreteness does not guarantee that users understand their underlying function (McDougall & Isherwood, 2000). It is therefore important to design IGOs whose link with significant events is easily understood by operators. The weight of evidence seems to point to high concreteness as a suitable characteristic for IGOs. This is supported by practice in TIP and ASI where the inserted objects are concrete representations of threats or faulty objects and cannot be distinguished visually from their real counterparts. However, it is possible that not all significant events can be depicted effectively in a concrete manner in IGOs.

Significant events on CCTV can be represented by a range of IGOs varying in their degree of concreteness. For example, an employee picking an item up from the top of a machine in an area where this is prohibited could be depicted by the concrete image of a person reaching for the item, thereby alerting operators to watch for this type of action. Alternatively, an abstract image could be used, consisting of an item with an arrow emerging from it to indicate movement away from its position, and a cross through it suggesting that movement is prohibited. An image with a combination of abstract and concrete features could include a person reaching for an object with a cross through it to suggest that this is prohibited. Research indicates that the meaning of the abstract image will be least accessible to users and would require most learning (Wogalter et al., 1997). However, this image may have more meaning for operators (with respect to the inferred behaviour; i.e. illegal removal of an object) than the image of a person with an arm outstretched.

Detection of significant events on CCTV relies on observation. This suggests that IGOs should depict observable objects, events or behaviours which indicate that a significant event is occurring in live video footage. An example of an important object could be a weapon or an abandoned bag. Behaviour could include a person swiveling his/her ankle to grind a diamond into the underside of his/her boot. An event may involve several people or a sequence of actions, such as individuals positioning themselves outside a building with strategic importance. The fact that these objects and events are observable suggests that they can be depicted in IGOs in a concrete manner. However, several challenges face designers attempting to create concrete IGOs. Some concrete images are limited in the extent to which they allow access to the referent depicted (McDougall & Isherwood, 2009) which is sometimes an inferred or (attempted) hidden behaviour.

The first challenge in designing concrete IGOs is that many significant events involve movement of some kind and this is difficult to represent in a comprehensible manner in a static image without resorting to abstract features such as arrows or textual labels. Secondly, designers must choose which subset of features to depict and how to do this, leaving room for interpretation (Familant & Detweiler, 1993). Thus even concrete images are abstractions of objects. These issues are illustrated by the example of an employee swiveling his/her ankle to grind a diamond into the underside of his/ her boot. This is difficult to express with a static image. Although the static image may be an accurate representation of a single moment in the swivelling process, it may not communicate the dynamic behaviour in a comprehensible manner, or in a way that makes the link between the IGO and its referent behaviour explicit.

A moving or dynamic IGO could model behaviour more effectively, but movement alters the conspicuity of the IGO (as discussed in section 5.6) and introduces additional technical issues in inserting the IGO into live video footage. Alternatively, a combination of concrete (e.g., leg, ankle and boot) and abstract (e.g., arrows depicting direction of ankle's movement) features may be more effective in communicating the behaviour. Because learning is required for abstract features (Forsythe, Mulhern, & Sawey, 2008), the introduction of abstract features is likely to have implications for training regarding the meaning of IGOs. This needs to be balanced with the finding that abstract images sometimes produce smaller semantic distances than concrete images (McDougall et al., 2001). Certain other behaviours may be easier to represent in a concrete manner, such as a person picking up an object.

While concreteness does not guarantee comprehension, it has the advantage of reducing the number of alternative interpretations (Dowse & Ehlers, 2004; McDougall et al., 2001). In addition, being able to name or recognise the graphic in a concrete image does not necessarily imply that the user understands its function (Isherwood et al., 2007). Similarly, detecting an IGO does not guarantee that the operator understands its link to significant events, in which case the IGO is unlikely to assist with the detection of significant events. This supports the need for small semantic distances, as discussed earlier.

Concreteness is clearly an important characteristic for IGOs. It is related to the understanding of images (McDougall et al., 1999, 2000, 2006). However, this effect is modulated by familiarity (McDougall & Isherwood, 2009; McDougall et al., 1999, 2009). The relationship between concreteness and understanding may be short-lived and disappear as users become more familiar with the image (McDougall et al., 2001). It is assumed that with IGOs, frequency of detecting the IGOs does not necessarily cause comprehension, once again emphasising the need for carefully designed IGOs that are easily understood. The characteristics of semantic distance, is likely to interact with concreteness in making the relationship between IGOs and their referent significant events more explicit and accessible.

In a study of safety signs, concrete pictorials were better understood than abstract ones but these effects could be overcome by training (Wogalter et al., 1997). Training increases familiarity and understanding and could alter the viewing process for operators who are trained on the meaning of IGOs. Concreteness is also related to the speed and accuracy of searching, but these effects decrease as users become more familiar with the icons (McDougall & Isherwood, 2009; Isherwood at al., 2007; McDougall et al., 2000, 2001). A review of the literature suggests that the most consistent contribution of concreteness is to the image's meaning and users' comprehension (e.g., McDougall et al., 2006).

5.3. Visual Complexity

Complexity is defined as the amount of detail or intricacy in an image (Snodgrass & Vanderwart, 1980, cited in Forsythe et al., 2008), but different types of details have been measured by researchers. Factors included in complexity measures are the number of shapes, letters and lines (García, Badre, & Stasko, 1994 cited in Isherwood et al., 2007), pieces of information (McDougall et al., 2006), presence and type of foreground (and background), number of objects and holes in them, edges and homogeneity in image structure (Forsythe, Sheehy, & Sawey, 2003), decomposability or the number of nameable and non-nameable visual parts (Lloyd-Jones & Nettlemill, 2007), and grouping of elements within images (McDougall et al., 2006). While all these details may be relevant to IGOs, they need to be reduced to a workable number. At an intuitive level, the most obvious details to consider for IGOs are the number of shapes, lines, objects and grouping of elements within IGOs.

A review of research on the complexity of visual stimuli has two key limitations in its applicability to the current research. Firstly, it has focused largely on its effects on the speed of search and to a lesser extent, ease of image interpretation. Studies on complexity and search speed examine searching for the objects themselves, rather than how perceiving the stimuli can assist the detection of other objects or events, as with IGOs. Therefore existing research tends to indicate how IGO complexity would influence the efficiency of detection of the IGOs themselves rather than how it could facilitate the detection of other significant events. The second limitation is that different results have been found for different contexts and search tasks. This suggests that the effects of complexity vary according to context and may not apply to CCTV. Despite these limitations, selected research is discussed in order to identify findings regarding complexity to consider in IGO design.

Existing research on complexity says little about how IGOs could assist in the detection of significant events, but may be useful in reducing one of the potential negative effects of IGOs – that of distracting operators from significant events. In certain contexts, complexity is negatively related to search efficiency, with simple visual stimuli being processed more quickly than complex ones (Byrne, 1993; Isherwood et al., 2007; McDougall et al., 2006; McDougall & Isherwood, 2009). Similarly, guidelines on images in other contexts, such as information and safety signs and icons in computer user interfaces, support the use of simple images (e.g., Byrne, 1993; Gittins, 1986). This is related to differences in the processing of simple and complex images which could have implications for the resources required to process IGOs and the length of time for which attention is removed from the rest of the display. Simple images are particularly useful in instances (e.g., hazard signs) that rely on their features to draw attention and therefore function on a largely stimulus-driven basis (Wickens & McCarley, 2008). Simple IGOs are likely to be conspicuous against their complex backgrounds and therefore may be detected easily. This would have the advantage of decreasing the time during which operators are distracted from potential significant events, but may not encourage an active search of the entire display.

Two explanations have been offered for the slower processing of complex images. The first draws on feature integration theory (Treisman & Gelade, 1980; Treisman & Sato, 1990) which proposes that it takes longer to bind multiple features in complex stimuli than the few features of simple ones (McDougall et al., 2006). This implies that practice or experience with complex images is unlikely to reduce search times because of the time required for visual attention processes (McDougall et al., 2006).

The second explanation focuses on the increased task demands posed by complex images and search which require greater working memory loads and result in slower response times (Shiffrin, 1988). This is consistent with vigilance research where signal complexity leads to decreases in overall levels of vigilance (Adams, Humes, & Stenson, 1962; Childs, 1976; Howell, Johnston, & Goldstein, 1966; Methot & Huitema, 1998; Montague, Webber, & Adams, 1965; Mouloua & Parasuraman, 1995). With ASI research, complexity is inversely related to inspection performance (Smith & Lucaccini, 1969). According to arousal theory, there is an optimal level of mental workload that, if exceeded, detracts from vigilance performance (Koelega, 1996). Complex stimuli increase mental workload and create a greater vigilance decrement (Sawin & Scerbo, 1995). This suggests that highly complex IGOs should be avoided to prevent potential negative effects on workload and vigilance performance. However, this conclusion is contradicted by research on other search tasks.

The finding that processing of complex images is slower than simple images is not consistent (Lloyd-Jones & Luckhurst, 2002). It applies mainly to synthetic stimuli in simple arrays (e.g., Treisman & Gelade, 1980; Wolfe, 1998) and icons in computer user interfaces (Li et al., 2005). However, different results have been obtained for different visual stimuli and tasks. Although the finding generally applies to synthetic stimuli, it is not supported for real objects and natural scenes (Li et al., 2005). The latter tend to be highly complex, yet scene categorisation appears effortless. This may be explained by the familiarity of natural scenes and the meaning imbued in them which guides search. Representations of stimuli in natural scenes exist in long term memory and these speed up search. This contrasts with many of the synthetic stimuli used in simple arrays which are unfamiliar and meaningless to observers and have to be learnt in order to be recognised. Similarly, the functions of icons in computer user interfaces need to be learnt and this could explain why some researchers have found that complex icons take longer to detect. Findings for natural scenes suggest that complex IGOs that have meaning for the observer may be processed quickly, overcoming the apparent disadvantages of complexity discussed earlier.

Based on the different findings in different contexts, it is evident that other factors, such as learning, familiarity, and meaning, influence the processing of stimuli, whether simple or complex. This also applies to the characteristics already discussed (semantic distance and concreteness). Furthermore, the characteristics interact with each other in visual stimuli. Therefore it is difficult to consider complexity in isolation. IGOs with small semantic distances, high concreteness, and high complexity may be processed differently from those with large semantic distances, low concreteness, and low complexity. Examples of the former include photographs or video clips of people performing prohibited actions, and the latter include simple, stylised images with symbols such as arrows. Depending on how these IGOs are positioned in the display, the latter may be more conspicuous and therefore more quickly detected. However, their meaning needs to be learnt and they may therefore have less impact on the detection of significant events. The former may be processed as quickly because they appear realistic, require less interpretation and have more immediate meaning. The finding that users tend to perceive images as being less complex over time (Forsythe et al., 2008) suggests that the effects of IGO complexity are likely to vary over time and any increases in search time may be overcome with practice.

Greater complexity or detail contributes to the realism of IGOs and their similarity to the significant events they are intended to represent. At an intuitive level, greater realism could assist in modelling the object, event or behaviour and consequently the interpretation of the IGO. Therefore it could be argued that greater complexity and realism are related to the

IGO's potential to develop and reinforce mental sets that could assist with the detection of significant events. This line of reasoning is consistent with initial research where it was believed that a high degree of complexity made images more realistic and easier to interpret (McDougall et al., 1999). However, complexity has recently been found to be unrelated to ease of image interpretation with concreteness being more important in this regard (McDougall & Isherwood, 2009). Therefore the above argument may not hold and although it requires consideration, complexity may be less important than other characteristics for IGOs. It is possible that relatively simple IGOs could portray the 'essence' of significant events equally or more effectively than complex ones.

Existing research provides few indicators regarding suitable levels of complexity for IGOs or how this could influence the detection of significant events. Complexity may affect the time taken to perceive and interpret IGOs. Therefore it could influence how accessible the meaning of IGOs is, and their ability to trigger mental models of significant events. However, based on McDougall and Isherwood (2009), it is possible that concreteness is more important than complexity in this regard. Different levels of complexity and the other characteristics combine in IGOs and are likely to have different effects on searching for both IGOs themselves and significant events. At an intuitive level, it seems that IGOs should contain sufficient complexity to portray the 'essence' of significant events. The amount of complexity needed is likely to vary for different significant events. However, conclusions regarding complexity are highly tentative as much is unknown and past research does not really provide a useful guide.

5.4. Salience

Research on salience provides little information about how IGOs could assist the detection of significant events. Rather, salience influences how conspicuous IGOs are and how easily they are detected. Perception of IGOs is a necessary 'first' step in facilitating the detection of significant events as IGOs have to be perceived before they can activate mental models of significant events. The perception of salient IGOs is likely to demand fewer resources than less salient IGOs, creating less demand on attention resources between IGOs and significant events. This is based on findings that less conspicuous events demand more resources and higher mental workload (Temple et al., 2000). Therefore salience has some relevance to the current research although its contribution may be assistance in minimising the potential negative effects of IGOs rather than direct assistance with detecting significant events.

Salience is defined as the conspicuity of visual stimuli (Wickens & McCarley, 2008), or the extent to which a feature or stimulus differs from its local surrounds and therefore ‘pops out' (Theeuwes, 1994). Salience in this context has a particular meaning which differs from its use in everyday English where it refers to importance. It also differs from semantic distance which does not consider the extent to which visual features of an image differ from the local surrounds.

It is emphasised that salience can only be assessed in relation to surrounding visual stimuli (McDougall et al., 1999). For example, a red stimulus amongst black stimuli is highly salient, but the same stimulus amongst other red or mixed stimuli, is less salient. Salience influences how quickly a stimulus is detected (Bauer, Jolicoeur, & Cowan, 1996b; D’Zmura. 1991) and its ability to draw attention (Itti & Koch, 2000). Therefore salience is likely to influence how quickly IGOs are detected and the extent to which they draw attention. This is important because it is assumed that IGOs should be detected quickly to reduce distraction from significant events that might be about to occur.

Past research has focused on the ability of salient stimuli to capture attention (e.g., Theeuwes, 1993; Treisman, 1988). A common finding is that salient stimuli tend to draw attention (e.g., Theeuwes, 1993; Treisman, 1988). However, as with complexity, findings differ between contexts and search tasks, particularly between simple stimulus arrays and real scenes, and passive or free-viewing and active search of scenes.

With simple stimulus arrays, singletons show a strong ability to pop out, draw attention and are detected very quickly (Theeuwes, 1993; Treisman, 1988; Treisman & Gelade, 1980; Treisman & Gormican, 1988). Singletons are stimuli containing one unique property, for example colour, luminance, form, orientation, or size and they may be static or dynamic (motion or onset) (Yantis, 1996). This is consistent with research on warning signs where simple images tend to stand out and draw attention (Wogalter, Godfrey, Fontenelle, Desaulniers, Rothstein, & Laughery, 1987). By contrast, search speeds for conjunctions (stimuli that differ from their surrounds on some features and share others)

(Itti & Koch, 2000) are more varied, and often slower than in simple stimulus arrays (Wolfe, 1998). The reasons for this are not clear, but may be related to greater salience, heterogeneity, and density or clutter (Berger & McLeod, 1996, cited in Wolfe 1998).

Search for triple conjunctions is more efficient than standard conjunctions of two features (Wolfe, 1998). This implies that search becomes easier and faster as more is known about the target. The search for categorical targets is fast, for example, searching for the image of a dog amongst letters of the alphabet. Here, the target and distracters are semantically different (Theeuwes, 1993). This may apply to IGOs which are a different category from the rest of the display showing real events. Examples include simple abstract and salient IGOs.

Research findings on salience in real scenes differ from simple stimulus arrays. The former research is usually based on photographs of real scenes (e.g., Einhaüser et al., 2008). Although these scenes are usually static, they are more similar to CCTV displays than simple stimulus arrays. Caution needs to be exercised in generalising results from laboratory research using simple stimuli against blank backgrounds to densely cluttered real world scenes (Bravo & Farid, 2006) including those in CCTV displays which are also dynamic.

Real scenes introduce greater complexity, clutter, noise, and cues regarding meaning to the scenes and impacts on the role of salience (Wogalter & Laughery, 1996). Salience plays a smaller role in gaze and attention control where people are searching real scenes in a purposeful manner than in simple stimulus arrays with little meaning (Henderson et al., 2007). Task goals, scene meaning, and memory for similar scenes and instances of observation play a dominant role in directing eye movements. Although salience plays a role in eye movements, areas which provide information relevant to the task are more important than salience in real scenes. This suggests that with CCTV, operators who are searching displays actively are likely to direct their gaze to areas that will provide information about significant events, rather than merely salient areas.

The ability of salient stimuli to draw attention is attributed largely to stimulus-driven attention processes (Serences et al., 2005). The relative influence of stimulus-driven and goal-directed attention guidance differs for simple stimulus arrays and real scenes. The ‘pop out' effect for salient singletons in simple arrays shows strong (but not exclusively) stimulus-driven processes (e.g., Theeuwes, 1992; Theeuwes & Godijn, 2002; Turatto & Galfano, 2000). However, with real scenes, goal-driven attention control dominates stimulus-driven control based on salient characteristics such as hue or luminance (Chen & Zelinsky, 2006; Ehinger & Brockmole, 2008; Einhaüser et al., 2008; Foulsham & Underwood, 2007, 2008; Henderson et al., 2007; Oliva, Torralba, Castelhano, & Henderson, 2003; Torralba, Oliva, Castelhano, & Henderson, 2006; Underwood, Foulsham, van Loon, & Underwood, 2005). It appears that salience has less power to draw the eyes in the search of scenes than it does in simple arrays (Henderson, 2007).

Explanations for the decreased ability of salience to draw the eyes in real scenes are based on heterogeneity and task demands. In heterogeneous scenes, salience is less conspicuous than in more homogeneous, simple arrays (Ehinger & Brockmole, 2008). Real scenes tend to be more heterogeneous, with each object being unique in some respect and this decreases the salience of individual items (Brockmole & Henderson, 2005b). With simple stimulus arrays, a unique distractor amongst otherwise similar stimuli tends to draw attention and can disrupt goal-driven search (Brockmole & Henderson, 2005b). However, when observers actively search a real scene for a complex target with low salience, high salience distractors have negligible effects on search and tend not to distract observers from their goals (Chen & Zelinsky, 2006; Foulsham & Underwood, 2008; Underwood et al., 2005). In this situation, high salience distractors can be overridden by task demands (Einhaüser et al., 2008; Underwood, Foulsham, van Loon, Humphreys, & Bloyce, 2006). This is likely to apply to CCTV displays which are heterogeneous and where operators have clear goals.

A number of factors moderate attention capture by salient stimuli. These include advance knowledge and expectations regarding stimulus features. This is supported by the singleton detection mode (Duncan, 1984), the contingent involuntary orienting hypothesis (Folk et al., 1992), inattentional (Mack & Rock, 1998), and change blindness studies (Rensink, 2002; Simons & Levin, 1997). Goal-directed processes can over-ride attention capture by the salient stimulus when there is advance knowledge of the target, such as its location (Yantis & Jonides, 2000) or colour (Folk et al., 1992; Wolfe, 1998). In this situation, salient stimuli that are irrelevant to the task do not affect performance (Folk & Remington, 1996). It is not clear whether these findings and theories apply to salient IGOs because IGOs are relevant to the operator's task. Operators are required to detect both IGOs and significant events and may therefore be unable to ignore salient IGOs. However, this may depend on how IGOs and significant events are prioritised by the operator.

One of the potential drawbacks of IGOs is that they could draw attention away from significant events when they occur simultaneously. However, salient stimuli do not draw attention when attention is already deployed elsewhere (Juola, Koshino, & Warner, 1995; Mortier, Donk, & Theeuwes, 2003; Müller & Rabbit, 1989; Simons, 2000; Theeuwes, 1991; Yantis & Jonides, 1990) or when they are not relevant to the task (Folk & Annett, 1994; Hillstrom & Yantis, 1994; Jonides & Yantis, 1988; Gibson & Jiang, 1998; Theeuwes, 1990; Todd & Kramer, 1994; Yantis & Egeth, 1999). This suggests tentatively that the potential for IGOs to distract operators from significant events as they unfold may not be significant. However, this may be related to how actively operators are observing versus passively 'watching'.

Active search involves goal-directed searching with deliberate eye movements and deployment of attention. With passive searching, fewer attention resources are likely to be deployed and scanning is likely to be more random. Highly salient stimuli have a greater ability to capture attention when there is spatial uncertainty and attention is spatially distributed or unfocused (Folk, Remington, & Johnston, 1992; Wolfe, 1998; Yantis & Jonides, 2000). Weaker endogenous processes are involved in these situations. This has two implications for the current research. Firstly, salient IGOs are likely to draw attention during passive search and this could have the advantage of directing attention back to the display. Secondly, when significant events are occurring and operators are observing them actively, salient IGOs are less likely to draw attention and distract operators from the significant event.

The above effects of significant events need to be weighed against the potential disadvantages of using IGOS that are consistently high in salience. Here, operators are likely to develop an attention set for highly salient IGOs that are easy to detect. While this may assist in reducing the duration of distraction from potential significant events, it could result in less active search involving processes similar to the singleton detection mode (Duncan, 1984). Here, observers search for stimuli with a feature that distinguishes them from the rest of the display, and reject other stimuli. Similarly, operators could search for salient IGOs and neglect the more difficult search for significant events. A set of IGOs with differing salience levels therefore seems more appropriate to prevent attention sets for IGOs from developing.

In addition to the study of real scenes, the role of salience in maximising the efficacy of warning and hazard signs and labels has been examined. These studies are useful in identifying features that contribute to salience and which should be considered in IGO design. Because signs and labels need to draw attention and be noticed easily (Wogalter & Laughery, 1996; Wogalter et al., 1997), they provide useful information on features that achieve this. Features that influence how easily pictograms are noticed include novelty, size, location, salient colour, intensity or luminance, contrast, signal words used, the nature of the pictorials, location (Wogalter & Laughery, 2006), and shape (Arend, Muthig, & Wandmacher, 1987, cited in McDougall et al., 2001). These features are comparable to those studied in the visual search and attention capture literature (e.g., Cavanagh, Arguin, & Treisman, 1990; Folk, Remington, & Wright, 1994; Hodsoll, Humphreys, & Braithwaite, 2006; Irwin et al., 2000) which also includes additional features such as motion (Rosenholtz, 1999), contour junction, edge termination, shading (Koch & Ullman, 1985), and orientation (Parkhurst, Law, & Niebur, 2002).

It is evident that many features contribute to salience. However, the features have different abilities to capture attention (Wolfe, 1998) and influence scan patterns (Findlay, 1997). Of the various features that contribute to salience, abrupt onsets exert the most robust form of attention capture (Brockmole & Henderson, 2005b). Abrupt onsets are stimuli that appear suddenly (Egeth & Yantis, 1997). This suggests tentatively that IGOs that appear suddenly are likely to be detected easily. Conversely, these IGOs may also have the most potential to distract operators from significant events. However, as with other salient stimuli, the ability to draw attention is modulated by goal-directed attention (Folk et al., 1994). It is possible that operators who are actively searching for significant events may not be distracted by salient IGOs at important moments in the search process.

The differential ability of salient features to draw attention is consistent with research indicating that there is a hierarchy of factors influencing scan patterns. Amongst bottom-up processes, it appears that salient colour exerts a greater influence than size or shape (Findlay, 1997). Foreknowledge of these features facilitates search. Motion and change predict eye movements better than colour, intensity and orientation (Itti, 2005). These features can be manipulated in IGOs to influence salience levels.

IGOs with high and low salience have both advantages and disadvantages. Past research does not provide clear indicators as to salience levels most appropriate to IGOs. However, it seems that a set of IGOs with mixed salience levels is likely to have the advantages of encouraging active search, reducing search processes similar to singleton detection mode, and reducing distraction from simultaneous significant events.

5.5. Novelty

Similarly to salience, novel stimuli tend to capture attention (Wickens & McCarley, 2008). Therefore novelty may impact on how quickly IGOs are detected and the level of distraction from potential significant events. The processes whereby this is seen as occurring are likely to be the same as for other salient characteristics and are therefore not discussed further here. However, novel IGOs with small semantic distances to significant events may also assist in developing mental models and attention sets for significant events that have not previously been encountered. Therefore certain novel IGOs may play a role in facilitating the detection of unfamiliar significant events.

Novelty and familiarity are closely related concepts which have been studied in the icon, visual search, and attention capture literature. Novelty is defined as ‘A new or unfamiliar thing or experience' (South African Concise Oxford Dictionary, 2002, p. 795). Familiarity denotes something that is well known, usual or common. Many authors tend to use these terms interchangeably (e.g., Wickens & McCarley, 2008). In the field of perception, novelty and familiarity refer to stimulus repetition or experience with stimuli (McPeek, Maljkovic, & Nakayama, 1999) and prior knowledge of stimuli (Suzuki & Cavanagh, 1995). Novel IGOs are IGOs that have not previously been encountered by the operator. IGOs could reflect images varying in familiarity to operators, such as abstract symbols or representations of everyday objects.

Familiarity is related to accessing meaning (Hancock & Rogers, 2004). Familiarity with images' meanings and not merely their appearance is important in searching as this is related to comprehension. Familiarity with the IGO's features and content may be insufficient to assist with the detection of significant events. An understanding of the link between the IGO and the significant event is needed for the IGO to assist with the detection of significant events. Exposure to novel IGOs may contribute to comprehension, but does not guarantee it. Examination of how novelty could be combined with other characteristics provides points as to how novel IGOs could alert operators to forms of significant events that have not previously been encountered.

When a novel object appears, the visual system creates a new object representation in a stimulus-driven manner (Enns, Yantis, & Di Lollo, 1998). Thus the new object captures attention through the creation of a new object file which triggers the redistribution of attention to the new object (Yantis & Hillstrom, 1994). Object files for IGOs would therefore need to be created for them to capture attention. Where semantic distances between IGOs and significant events are small, the object files for IGOs are seen as reinforcing those for significant events. In this way, novel IGOs could assist with creating mental models and attention sets for significant events that have not previously been encountered.

Perception of novel IGOs and the creation of new object files are not seen as being sufficient for the former to facilitate the detection of significant events. The link between IGOs and significant events also needs to be understood by operators if IGOs are to influence attention sets for significant events and guide attention in a goal-driven manner. Novel IGOs with small semantic distances and which are highly concrete may be more easily understood than novel IGOs with large semantic distances and/or abstract images. As discussed previously, concreteness is related to comprehension (McDougall et al., 1999, 2000, 2006). This suggests that novelty should be combined with small semantic distances and high concreteness for these IGOs to be effective.

Novelty by itself is insufficient to capture attention (Gellatly, Cole, & Blurton, 1999). New objects that are low in salience do not always capture attention. This suggests that novel IGOs should also be salient if they are to be detected quickly. Therefore it is recommended that novel IGOs also have small semantic distances and high concreteness and salience.

5.6. Realism

Realism influences the extent to which IGOs blend in with the scene in terms of both physical features and the scene's gist. Earlier research reported that objects that are incongruent with the gist of scenes are detected quickly (Davenport & Potter, 2004; Gordon, 2004). However, other studies have found that this is not the case (De Graef, Christiaens, & d'Ydewalle, 1990; Henderson, Weeks, & Hollingworth, 1999). A recent study reported that incongruent objects are only detected more quickly than congruent objects when they are also visually salient (Underwood & Foulsham, 2006). Realism is seen as being important to IGOs as it may influence how accurately the relationships between IGOs and their referents are understood, and in turn the impact of IGOs on the detection of significant events. In addition, CCTV displays show realistic scenes and therefore realism is likely to influence IGO salience.

All characteristics discussed above are likely to influence the degree to which IGOs appear realistic. Complexity was originally perceived as an important contributor to realism (McDougall et al., 1999), but hue, luminance, concreteness, semantic distance, and dynamic properties are all likely to be important. Abstractness and semantic distance influence realism in that they are likely to impact on the degree to which IGOs are congruent with the physical features and context or meaning of the scene, or how likely it is for the graphic depicted to be part of the scene.

Additional features such as size, orientation, angle, perspective, shading, shadows, gloss, and the location of IGOs need to be considered if realistic IGOs are needed. Given the numerous factors that affect realism, it is apparent that highly realistic IGOs are difficult to design as they are very specific to camera views and how the camera is operated, such as zooming in on objects. Highly realistic IGOs which fit the scene in all respects (e.g., size, depth of field) are likely to be less salient than less realistic IGOs.

The least salient and most realistic IGOs are likely to be those consisting of real or enacted footage which is placed in scenes with the same context and 'blends' into displays. Although these may encourage active search, they are likely to suffer certain drawbacks. They may be large and could obscure certain camera views, preventing the detection of actual significant events that occur while the inserted footage is displayed (Donald & Donald, 2008). They would need to be inserted in a section of the display where the context is appropriate and aspects such as perspective, size, colour and depth of field are appropriate to the display, increasing technical challenges. This is further complicated by cameras that can pan, tilt or zoom which changes the contextual display. In addition, operators are likely to know about significant events that have occurred, (even if they did not detect the significant events themselves) because significant events may be discussed amongst surveillance staff or used in training sessions. A large pool of such IGOs would need to be developed to reduce familiarity.

The realism of TIP images in X-ray screening is likely to differ from IGOs in CCTV. TIP images are highly realistic and replicate actual threats but a similar degree of realism is difficult to achieve with CCTV due to the dynamic properties of CCTV displays, the operator's ability to control cameras, changing viewing conditions, the use of multiple displays, and the time taken for significant events to unfold (Andrew et al., 2003). In conclusion, it would be difficult to design IGOs that are highly realistic in appearance and are inserted in a way that maintains their realism. A highly sophisticated technical system could possibly achieve this, but no such system currently exists. Static photographs and dynamic video clips currently provide the most realistic material for IGOs, but the realism may decrease once they are inserted into live video footage where numerous factors need to be considered to make them blend into the scene in a congruent manner.

Realistic IGOs with small semantic distances are highly concrete, complex, and could include varying degrees of novelty. They have the advantage of depicting the 'essence' of significant events in a direct manner that should be relatively easily understood by operators. However, realism may not be a necessary characteristic for IGOs as less realistic images may also be able to convey observable and important aspects of significant events. It is possible that IGOs with low to moderate realism may be effective in replicating significant events when combined with small semantic distances and sufficient detail to depict the significant events they represent.

5.7. Categorisation of Characteristics

A number of characteristics have been covered in the literature review above. This section discusses how they could be grouped at a conceptual level, rather than how they could be combined physically in IGOs, which is discussed in section 5.8. Salience and realism are the most inclusive characteristics in that they include numerous features. Factors that have typically been considered under salience are hue, luminance, novelty through new objects, dynamic properties in the form of abrupt onsets, and basic features such as size, form, edge orientation, and gloss (Koch & Ullman, 1985; Parkhurst et al., 2002; Rosenholtz, 1999). Some researchers have grouped these features under salience, while others have treated them as separate characteristics. For example, novelty is considered part of salience (Butter, 1987), but is also listed as a separate characteristic (e.g., Wickens & McCarley, 2008).

In the current research, salience includes hue, luminance, dynamic properties, novelty and the basic features of size, orientation and gloss. These characteristics are most likely to contribute to salience in IGOs. As previously stated, all characteristics contribute to realism, including salience. However, it does not make sense to categorise all characteristics under realism. Realism appears to result from the combination of the other characteristics as well as how IGOs are inserted into the context of dynamic displays.

Research in different areas has tended to focus on different characteristics with the result that there are no typologies that include all characteristics covered in this chapter. Concreteness, complexity, semantic distance, and realism have been examined in literature on icons and pictorials rather than visual search and attention capture. The latter have adopted a strong focus on salience. Characteristics which have not traditionally been included in studies of salience are nevertheless likely to modulate salience in CCTV. These are concreteness, complexity, semantic distance, and realism which affect the degree to which an image is congruent with its context. The way in which these characteristics draw attention may, however, differ from the characteristics traditionally included in salience studies in the visual search literature.

Salience is typically reported to function on a largely stimulus-driven basis, with modulation by goal-directed attention processes (e.g., Mazer & Gallant, 2003; Remington, Folk, & McLean, 2001). Therefore the detection of salient characteristics also involves goal-driven attention. With CCTV, goal-driven attention guidance is likely to be based on attention sets for significant events and IGOs. Congruence between attention sets, visual stimuli, and the display context is crucial to this process. Characteristics that are most likely to influence congruence in terms of meaning (rather than physical properties such as luminance or hue) are those that have not traditionally been examined in the visual search and attention capture literature - concreteness, complexity, semantic distance, and realism. The salience of these characteristics depends on their meaning rather than merely physical properties. Therefore it seems appropriate to categorise these as separate characteristics.

Table 3 indicates how the characteristics are categorised conceptually in the current study.

Table 3. Conceptual categorisation of IGO characteristics

Characteristic	**Facilitate IGO and/or significant event detection**
Semantic distance	Significant event
Concreteness	Significant event
Complexity Lines Objects Grouping	IGO
Salience Hue Luminance Abrupt onset Onset of movement Size Form Orientation Gloss	IGO
Novelty	Significant event and IGO
Realism	Significant event and IGO

Six characteristics have been identified. Although novelty contributes to salience, it is seen as contributing to the detection of both significant events and IGOs and is therefore included as a separate category. It is apparent that only four characteristics are likely to assist the detection of significant events. The remaining characteristics influence the detection of IGOs but may not facilitate the detection of significant events apart from drawing attention to IGOs that depict these events. The way in which the characteristics are combined in IGOs is seen as affecting search and their relative importance, as discussed next.

5.8. Interactions Between IGO Characteristics

Literature regarding images indicates which characteristics may be appropriate for IGOs, these have previously been considered in isolation from each other and no guiding theory regarding combined characteristics exists. Such theory would result in more purposeful or optimised rather than relatively 'randomly' designed IGOs. Research into complex monitoring requires a departure from the traditional approach of studying vigilance by isolating and manipulating simple variables. This is not possible in complex monitoring due to the very nature of complex displays where multiple facets are inter-related (Howell et al., 1966). Because detection involves both perceptual and cognitive processes based on inter-related aspects of the display, the composite impact of the overall display is more important than the effects of discrete characteristics or stimuli themselves. Due to their nature, complex IGOs will combine a number of characteristics. Therefore the composite impact of the characteristics will be explored and evaluated rather than the distinct and separate characteristics themselves.

Different physical combinations of characteristics could result in an extremely large range of possible IGOs. Therefore it is important to consider the relative importance of the characteristics and which combinations are most effective in enhancing the detection of significant events. Based on the literature reviewed in this chapter, certain characteristics are recommended for IGOs which reduce the range of IGOs seen as being appropriate for CCTV. These are small semantic distances, high concreteness, variations in salience (including hue, luminance, and size), novelty and realism, and low to moderate complexity.

It is proposed that IGOs representing important and observable aspects of these events (i.e., small semantic distances to significant events) assist in developing, reinforcing and activating mental schemas, models, and attention sets for these events. This process is supported by research indicating that images can assist in the development and activation of mental schemas for events, functions and processes (McDougall et al., 2001). Schemas and mental models form part of SA which is important in the detection of significant events (Donald & Donald, 2008).

Goal driven processes modulate the ability of stimuli to capture attention, search speed and accuracy. Key amongst these is consistency with the attention set (Wickens & McCarley, 2008). Therefore IGOs with small semantic distances to the significant events they represent are seen as requiring similar attention sets to significant events and are likely to be detected more quickly than those that are dissimilar. Research using real world scenes indicates that task demands are more important than salience in controlling attention (Einhaüser et al., 2008; Underwood et al., 2005). Thus top-down processes such as SA, attention sets, and task demands are likely to be very important in the detection of IGOs and significant events in CCTV which records real scenes.

The use of IGOs with small semantic distances to significant events is seen as assisting in the creation of expectancies and attention sets for significant events. Expectancies and attention sets play an important role in visual search and modulate distraction by stimuli that are not relevant to the task (Simons & Chabris, 1999). They contribute to goal-directed attention guidance (Chen & Zelinsky, 2006) and therefore could assist operators in searching for objects and events in the display that could signal significant events. Activated schemas and attention sets relevant to significant events are part of goal-directed attention guidance and therefore assist selective attention by biasing the attention filter used to reduce the number of stimuli for processing (Wickens & McCarley, 2008).

Semantic distance is considered the most important IGO characteristic in influencing goal-directed attention processes through the use of mental schemas, expectancies and attention sets. This is supported by its role in image processing in other contexts such as safety signs and icons in computer user interfaces (e.g., McDougall et al., 1999; Wiedenbeck, 1999). It is also likely to contribute to operators' understanding of the significant object or event represented by the IGO. This is consistent with the finding that semantic distance and familiarity are the most important characteristics in identifying the correct icon in computer user interfaces (McDougall & Isherwood, 2009).

In addition to semantic distance, concreteness is seen as important in guiding attention on a goal-driven basis to significant events, especially where familiarity with IGOs is low. Concrete IGOs could assist goal-driven search by providing a direct representation of the object or event involved in significant events. This is supported by the importance of both semantic distance and concreteness in icons in computer user interfaces (Isherwood et al., 2007). However, some abstract images also have small semantic distances (McDougall & Isherwood, 2009) and therefore have the potential to be more effective than certain concrete images. This may apply to the CCTV context, where important aspects of significant events, such as movement, are difficult to convey with static images. Therefore it is recommended that concrete IGOs are used wherever possible but that abstract features are included where necessary. However, operators may need training on the meaning of IGOs with abstract features.

At an intuitive level, complexity may contribute to an understanding of the meaning of images (Lloyd-Jones & Luckhurst, 2002) or the object or event that IGOs depict. However, research indicates that complexity does not necessarily contribute to the recognition of images (McDougall et al., 2006). In addition, guidelines for images in various contexts generally recommend simple images (Gittins, 1986). This suggests that IGOs should be simple, but that sufficient detail should be included to convey the meaning of the image. If users understand what significant object or event is depicted in the IGO, the latter could assist with goal-driven search. For some significant events and objects, a certain amount of detail may be required for users to comprehend their link with significant events.

In addition to goal-directed attention, stimulus-driven attention control is involved in the search process although its effects are not as strong in real scenes as in simple stimulus arrays (Foulsham & Underwood, 2008; Henderson et al., 2007). Several explanations have been provided for this. First, real scenes are imbued with meaning which interacts with salience in guiding attention (Li et al., 2005). Second, task demands are important in guiding search in real scenes (Einhäuser, Rutishasuer, & Koch, 2008). When observers are free-viewing (i.e., looking at a scene without a task goal) a scene with no task demands, salience guides attention more strongly. Operators' searching processes should be guided by the tasks of detecting significant events and IGOs. However, there may be periods of time when they free-view displays rather than actively searching them, in which case salient IGOs may draw attention more strongly. Last, real scenes are heterogeneous and this influences the salience of certain stimuli. As CCTV displays contain real scenes, the ability of salient IGOs to capture attention due to their salience is likely to be comparable to other real scenes used in other research. Therefore salient IGOs may not draw attention as strongly as salient stimuli in simple arrays of meaningless stimuli. Nevertheless, salience is likely to be an important contributor to stimulus-driven attention guidance and to play a role in search.

When IGOs combine high salience and small semantic distances, the stimulus-driven effects of salience and goal-directed guidance of small semantic distances and their related attention sets are likely to interact in directing attention. Therefore the combination of salient IGOs that represent significant objects or events are likely to utilise both top-down and bottom-up attention guidance to assist IGO and the detection of significant events. When salient IGOs draw attention, the content of the IGOs with small semantic distances should reinforce attention sets for significant events. Researchers acknowledge that goal-directed attention processes modulate stimulus-driven processes (Serences et al., 2005). At an intuitive level, it makes sense for IGOs to reflect characteristics that assist both forms of attention guidance in supporting each other rather than working against each other. The combination of semantic distance and salience is seen as key to this process.

Novelty and familiarity, combined with salience, influence attention capture (Gellatly, Cole, & Blurton, 1999). Novel IGOs are most likely to draw attention when they appear suddenly in the display, represent a luminance increment, and have small semantic distances. This combines stimulus-driven characteristics associated with efficient search and goal-driven processes where IGO content is aligned with attention sets for significant events.

Research regarding the interaction and relative importance of image characteristics in other contexts has tended to focus on semantic distance, concreteness, complexity, and familiarity of icons used in computer user interfaces (e.g., McDougall & Isherwood, 2009; Wiedenbeck, 1999). McDougall and colleagues are amongst the few researchers who have compared the relative contributions of most characteristics (e.g., Isherwood, McDougall, & Curry, 2007; McDougall & Isherwood, 2009). In addition, the CCTV context and task has not been included in previous studies on search or images. Little is known about search and image characteristics in CCTV. Therefore it cannot be assumed that past research in these other contexts applies necessarily to CCTV. The current study contributes to addressing this shortfall.

Different contexts use different types of visual stimuli (e.g., types of images and backgrounds) and search tasks. The contexts examined in this chapter are the search for icons to activate computer functions, noticing and interpreting warning signs on objects or in public spaces, and detecting IGOs and significant events in CCTV. Semantic distance is important in all three contexts mentioned above. There are indications that it is the most important characteristic for icons in computer user interfaces when users are not familiar with the icons (McDougall & Isherwood, 2009). This may be due to the need for understanding of the relationship between the icon and its function in order to identify and use it effectively (McDougall & Isherwood, 2009). Similarly, semantic distance is very important in warning signs (Mansoor & Dowse, 2007). Here, salience is also extremely important in order to attract attention to the signs (Wogalter & Laughery, 1996). This is similar to CCTV where IGOs should meet both requirements of activating attention sets relevant to significant events and attracting attention quickly. Therefore it is proposed that semantic distance and salience are most important to IGOs.

In addition to semantic distance and salience, concreteness is important in certain contexts. Concreteness is important when users need to understand images immediately and are not familiar with them (Isherwood et al., 2007). This applies to computer user interfaces and warning signs (Wogalter & Laughery, 1996) and is likely to be relevant to IGOs. However, not all functions and events can be depicted with concrete images in a way where the meaning is accessible to all users (McDougall & Isherwood, 2009). Therefore abstract images are also used. Although abstract images take longer to learn and understand, they have the advantage of representing some events using a smaller semantic distance than a concrete image would (McDougall et al., 2001). With CCTV, some IGOs need to incorporate abstract features in order to depict significant events in a comprehensible manner. With computer user interfaces, semantic distance is more important than concreteness (McDougall et al., 2001). Once again, it seems logical that this applies to CCTV due to the need for IGOs to reinforce attention sets for significant events.

There is likely to be strong relationship between complexity and realism (McDougall et al., 1999). However, complexity and realism are not the same characteristic as numerous characteristics contribute to realism. Guidelines generally recommend simple images (e.g., Byrne, 1993; Gittins, 1986), suggesting that realism is not a key characteristic in detecting and understanding images. It is recommended that IGOs are simple, but contain sufficient detail to convey their link with significant events.

In addition to contextual factors, the relative importance of image characteristics is influenced by individual and organisational factors and information processing stages. Users' familiarity with images influences the role of semantic distance in computer user interfaces, with familiarity becoming more important once learning of the images has levelled off (Isherwood et al., 2007; McDougall et al., 2001). With CCTV, familiarity with the IGOs and an understanding of the significant events they represent are likely to have similar effects, with semantic distance being most important initially and then familiarity making the greatest contribution when IGOs have been learnt.

Based on the changing relative contributions of characteristics to search efficiency and accuracy, it is unlikely that a single characteristic makes the greatest contribution in all contexts, tasks and across all stages of information processing. However, the importance of semantic distance has emerged fairly consistently in the literature reviewed, suggesting that it forms an appropriate starting point for research into IGO characteristics.

5.9. Characteristics Examined in Current Research

From a critical review of the theory and empirical evidence, only four characteristics are seen as influencing the detection of significant events – semantic distance, concreteness, novelty and realism. The current research focuses specifically on semantic distance because this is arguably the most important characteristic in developing and maintaining mental models and attention sets that are relevant to significant events. IGOs with small semantic distances to significant events are likely to assist in the creation and maintenance of attention sets for significant events that occur in a similar manner to those depicted in these IGOs. This contrasts with IGOs with large semantic distances to significant events or IGOs which are not in any way related to significant events, and which are therefore unlikely to facilitate the creation and maintenance of attention sets that are relevant to significant events. In addition, previous studies looking at IGOs (Andrew et al., 2003; Neil et al., 2007) have used a variety of images with varying semantic distances and random images, but have not examined the effects of semantic distance or the types of images used. The former study used a set of IGOs with mixed semantic distances and the latter used random IGOs without referents. This could have impacted on the results of the studies. The current study differs by including three groups with different treatments – no IGOs, random IGOs (RIGOs) and IGOs with small semantic distances to significant events (SIGOs). It is noted that the term 'random' as applied to IGOs in this research is used with a specific meaning. It does not imply a random selection of IGO content with varying semantic distances to significant events. Rather, it only refers to IGO content that is not intended to be related to significant events in any way. This differs from arbitrary semantic distances where there is a relationship between the image and its referent although this has to be learnt (Isherwood et al., 2007). With RIGOs, there is no relationship to significant events and no need for learning.

The likely effects of RIGOs on vigilance performance are unclear, with different theories making different predictions. Expectancy theory and SDT (MacMillan & Creelman, 1991) predict that the increased frequency of events requiring a response will alter the response criterion, resulting in more frequent hits (although this is more frequent hits on the RIGO, not more frequent hits on the significant event). This is supported by vigilance research (See Howe, Warm, & Dember, 1995), provided that the event rate is not too high (Davies & Parasuraman, 1982). However, multiple resource attention theory indicates that frequent stimuli may place too large a demand on visual attention resources and result in resource depletion and decreased hits (Smit et al., 2004; Wickens, 2002). In addition, the earlier review of literature (e.g, Wickens & McCarley, 2008) indicates that RIGOs could distract operators from significant events, resulting in decreased detection of significant events.

Concreteness is the second characteristic that is seen as influencing the detection of significant events. It affects the way in which the significant event is depicted in the IGO. Varying levels of concreteness can be associated with small semantic distances. An IGO with a small semantic distance can include abstract or concrete images, or combinations of these. In the current research all SIGOs were highly concrete to reduce the time required for learning and interpretation. This has the added benefit of holding concreteness constant in order to avoid conflation. RIGOs included objects and shapes (e.g., circle, leaf) that varied in concreteness but which are commonly used in various contexts of daily life. Participants were not required to name the shape or object, but only to report that it was perceived. Therefore it did not matter if participants did not know the name of the object as they merely had to perceive it.

Novelty is the third characteristic proposed as assisting in the development mental models of significant events. As none of the participants have been exposed to IGOs previously, they were not familiar with the IGOs. In addition, the context of the videos shown is a diamond processing plant. None of the operators had worked in a diamond processing plant. Therefore the nature of the significant events was novel to them even though there may have been some similarities to their current work contexts. Therefore it is assumed that novelty is held fairly constant for participants.

The remaining characteristics are seen as influencing the detection of the IGOs themselves. Although this is important in ensuring that IGOs are perceived and therefore have the chance to influence mental models and significant events, they are beyond the scope of the current research. However, they need to be considered in the design of the IGOs as they are all present in IGOs. Because IGOs consist of combinations of characteristics rather than single characteristics, it is not possible to design IGOs with only one characteristic. All IGOs have certain levels of concreteness, salience, and other characteristics. This is elaborated upon in the methodology chapter. It is recommended that future research examine the contribution of other characteristics and map the interactions and relative contributions of other characteristics.

5.10. Hypotheses

In the current research it is important to identify detection rates in order to establish the overall level of vigilance performance and whether a vigilance decrement occurs over time. These indicate whether there is the opportunity for the IGO intervention to enhance detection levels.

Previous controlled CCTV studies with simple known tasks have produced detection rates of about 90% (van Voorthuijsen et al., 2005; Wallace et al., n.d.). In the current research, all TBs are visible, no inferences are required, and the easiest and most difficult TBs were removed during the pilot studies. However, in more complex work settings involving CCTV surveillance, such as Andrew et al. (2003), appreciably lower detection rates were obtained. Further, Donald (1998) indicated that there were large differences in detection rates between individuals in a diamond production plant. Research into CCTV success has also indicated that the viewing approaches adopted by operators vary and can have a major impact on the success CCTV outcomes (Wells et al., 2006). A review of vigilance research using laboratory experiments indicates that actual rates of detection are seldom reported. In the current research, detection rates set the baseline of the overall level of vigilance performance and whether there is the opportunity for the IGO intervention to enhance detection levels. Given the greater complexity of the realistic environment, the more discrete nature of the TBs to be detected, and the greater requirements for visual analysis, the first hypothesis is as follows:

Hypothesis 1: The rate and accuracy of detection of realistic target behaviours (TBs) in a complex CCTV surveillance work environment is lower than for known and visually salient significant events in controlled CCTV settings.

This hypothesis allows for the establishment of detection rates and accuracy, and comparison of these with existing CCTV research. Detection is measured in terms of target behaviour detection rates, false alarms (FAs) and overall detection accuracy which consists of a combination of normalised TB and FA scores.

The second hypothesis is aimed at establishing whether a vigilance decrement occurs on the CCTV surveillance task used in this research. Most vigilance research has demonstrated a decrease in detection rates and accuracy over time (Craig, 1985; Hollenbeck, Ilgen, Tuttle, & Sego, 1995; Mackworth, 1957; Parasuraman, 1979, 1986; Molloy & Parasuraman, 1996; Parasuraman, Warm, & Dember, 1987; Pigeau, Angus, O'Neill, & Mack, 1995; Schmidtke & Micko, 1964; Warm, 1984; Wiener, 1987). In addition, the vigilance decrement is associated with tasks with a high mental workload and stress (Temple, Warm, Dember, Jones, LaGrange & Matthews, 2000; Warm, Parasuraman, & Matthews, 2008), decreased signal conspicuity (Becker, Warm, Dember, & Hancock, 1991), high event rates, spatial and display uncertainty, noise (Parasuraman & Mouloua, 1996), frustration (Warm, Dember, & Hancock, 1996), and successive discrimination tasks (Parasuraman & Davies, 1977) where observers are required to remember what significant events look like. These conditions are present in CCTV surveillance and therefore the vigilance decrement is likely to occur. The second hypothesis is:

Hypothesis 2: There is a vigilance decrement over time on the designated ninety-minute task.

If present, statistically significant decrements will be present in TB detection rates, FAs reported, accuracy and IGOs over the thirty-minute phases.

The third hypothesis examines the effectiveness of the IGO intervention in enhancing the detection of TBs. SIGOs and RIGOs are the two broad types of IGOs used in the current research. SIGOs have small semantic distances to their referent significant events rather than the scene context. They are aimed at enhancing the detection of significant events by developing and reinforcing relevant SA, mental models, and attention sets. This should assist goal-directed search and attention guidance. Small semantic distances are seen as facilitating this through the close link between SIGOs and significant events. The shapes and objects depicted in RIGOs, on the other hand, are not related to significant events, reflecting random semantic distances as explained in sections 5.1 and 5.9. None of the RIGOs represented significant events in any way. Therefore SIGOs are proposed as having a greater ability than RIGOs to facilitate the detection of significant events. This gives rise to the third hypothesis:

Hypothesis 3: IGOs enhance detection performance over time, with SIGOs being more effective than RIGOs.

The next chapter explains the methodology, materials, and vigilance task used in the research.

CHAPTER 6: METHODOLOGY

This chapter describes the research design, sample, procedure, the pilot studies involved in developing the materials for the research, analysis methods, and ethical issues involved. The aims of the study were to evaluate the detection performance levels obtained on the designated vigilance task, to establish whether a vigilance decrement occurred over the ninety-minute period, and to evaluate the effects of the intervention on detection performance. The background to the intervention and hypotheses were outlined in chapters 3 and 5.

A matched three-group quasi-experimental design was used, with groups referred to as the benchmark, RIGO, and SIGO groups. All groups observed the same video, but received different treatments. The benchmark group observed the video without any IGOs. The RIGO and SIGO groups had random object IGOs and IGOs with small semantic distances to significant events inserted into the video respectively. The benchmark group was required to detect significant events, and the treatment groups to detect both significant events and IGOs. This design allowed for comparison of performance between the groups, identification of potential vigilance decrements, and identification of the effects of different types of IGOs on the detection of significant events.

The research took place in two stages. Stage 1 included a sample of 42 operators who were employed full-time in CCTV surveillance. In view of unexpected results regarding potential vigilance decrements and high variance in target behaviours (TBs) and false alarms (FAs), it was decided to extend the sample to include novices, or participants with no experience with CCTV surveillance. The rationale for this is explained in more detail in section 7.1.3. Stage 2 included the original operators as well as 31 students who were novices regarding CCTV surveillance. The total sample of 73 was divided into three strata, with operators categorised as generalists and specialists, and students as novices. The terms 'generalist' and 'specialist' referred to surveillance background which consisted of the type of surveillance tasks performed at the organisations where they worked. This is explained in more detail in sections 7.1.2 and 7.1.3. The same research design, procedures, materials and types of analyses were used for the two stages. Exceptions to this are mentioned where relevant.

6.1. Samples

This section describes the samples for the pilot studies and Stage 1 only, as the additional sample for Stage 2 is described in chapter 7 with the results.

6.1.1. Pilot studies samples.

The sample for pilot study 1 on clips with TBs consisted of six people. Participants were faculty members who participated voluntarily, aged between 25 and 58. For the non-TB clips, participants were three people with expertise in detecting TBs in the diamond processing industry.

6.1.2. Stage 1 sample.

Sixty eight professional operators participated in the sample for Stage 1. Non-probability purposive sampling was used (Rosenthal & Rosnow, 2008) and participation in the research was voluntary. Participants all performed CCTV surveillance in their jobs although the degree of live surveillance required varied, with some performing live surveillance in real time, others reviewing past footage and others doing a combination of these forms of surveillance in their jobs. All had normal or corrected to normal vision. Participants were drawn from six organisations in the hospitality and gaming, town centre surveillance, academic, mining, and processing sectors. The number of participants from each organisation ranged from seven to seventeen, with a mean of twelve. Examples of job titles are: surveillance officer, control room operator, review and evidence specialist, surveillance supervisor and senior security officer. No participant had worked in a diamond processing plant previously and none had an advantage in terms of SA for the video used in the study.

On administering SAMAE (described in section 6.2.1) and the video (described in section 6.2.2), it was found that a large proportion (N=26, 38%) of participants obtained excessively high FA scores on SAMAE and/or the video. Seventeen participants reported extremely high FAs on SAMAE and sixteen on the video (Table 4). Examination of individual responses indicated that these participants appeared to be ‘guessing’ or responding randomly. This is similar to response sets encountered in questionnaires (Messick, 1962). Three participants were excluded for other reasons such as falling asleep on the task and responding inappropriately. Therefore these participants were eliminated as their responses did not contribute to an understanding of the detection process over time. The elimination of these participants still allowed for a range of detection scores and FAs on SAMAE and the video, but excluded those with FAs that were so high as to render the results meaningless. The final sample consisted of 42 useable respondents. This included a range of three to fourteen operators from each of the six organisations (mean=7).

Table 4. Reasons for excluding operators

Reason for exclusion	N	% of excluded operators
SAMAE FAs	17	65.38
Video FAs	16	61.54
Other	3	11.54

There were considerably more men (32) than women (10) in the sample (Table 5).

Table 5. Gender and education by experimental group (N=42)

	Group			
	Video only (n=13)	**RIGO** (n=14)	**SIGO** (n=15)	**Total** (N=42)
Gender:				
Male	11	11	10	32
Female	2	3	5	10
Education:				
< Grade 12	1	1	4	6
Grade 12	12	11	7	30
> Grade 12	0	2	4	6
Age: Mean (SD)	38.07 (9.11)	36.2 (6.91)	35.08 (8.89)	36.48 (8.2)

Participants were allocated to experimental groups on the basis of SAMAE scores (described in section 6.2.1). This resulted in two women being in the non-treatment or benchmark group, three in the RIGO and five in the SIGO group. Participants' ages ranged from 21 to 53, with a mean of 36.48 years and standard deviation of 8.2 years. The mean age for each treatment group was similar. The modal education level was grade 12 (n=31). Three participants had between grades 8 and 11, seven had obtained certificates or diplomas in addition to grade 12, and one person held a Bachelor's degree. Twenty four participants were African, sixteen were White and two were 'Coloured' – based on official South African Government classifications in the Employment Equity Act of 2004. These characteristics were spread fairly evenly between the treatment groups.

Regarding the CCTV surveillance experience of the sample, almost two-thirds (N=26, 62%) had five or fewer years' experience with CCTV surveillance, and only a few (N=4, 9.52%) had more than ten years' experience (Table 6). The type of surveillance performed varied, with almost a third of participants performing live surveillance only (n=12, 29%), only three (7%) reviewing historical footage only, and over half the sample (n=27, 64%) performing both live surveillance and reviewing. Most participants (n=36, 86%) had received training in CCTV surveillance and observation skills, although six (14%) had only received on-the-job or other types of technical CCTV training. The majority of participants were in the Paterson B-upper (38%) and C-lower grades (52%) and the remaining ten percent were in the C-upper grade. Participants in the experimental groups were regarded as having comparable experience, training and roles in CCTV surveillance.

Table 6. CCTV surveillance experience, type of surveillance, training and job grade by group (N=42)

	Group			
Background	**Video only** (n=14)	**RIGO** (n=15)	**SIGO** (n=13)	**Total** (N=42)
Experience				
≤ 5 years	10	9	7	26
6-10 years	3	4	4	11
≥ 11 years	1	1	2	4
Unknown	0	1	0	1
Surveillance type				
Live only	5	5	2	12
Review only	0	2	1	3
Both	9	8	10	27
Training				
Technical	1	2	3	6
Surveillance	6	6	5	17
Combination	7	7	5	19
Grade				
B-upper	5	6	5	16
C-lower	7	9	6	22
C-upper	2	0	2	4

6.2. Procedure

The same procedures were used for Stages 1 and 2. Any differences are noted in the results chapter for Stage 2. Access to the organisations was requested via an email, phone call and letter (Appendix 1) to the security or asset protection managers of relevant organisations. Once permission had been obtained to conduct the study, security managers distributed the participant information letter (Appendix 2) to CCTV surveillance personnel. Participation was voluntary and was confirmed by each participant signing a declaration to this effect (Appendix 3). The involvement was arranged at a time convenient to participants and organisations. The organisations used a variety of shift systems but all arranged for research to be conducted on the participants' off days. Participants were compensated for taking part during their day off by being given an additional day's leave or by being paid for their time by the organisations. Therefore participants were not disadvantaged by participating. These measures were initiated by the participating organisations and were the standard procedures used for activities such as training which typically do not take place during routine work hours.

Participants completed SAMAE (Donald, 1998) in order to provide a basis for allocating them to treatment groups (explained below). They were then given a thirty-minute break, after which they observed the video. SAMAE and the video together yielded a time of about 4 hours per participant. Although it would have been preferable to administer SAMAE on a separate day from the video, this presented logistical difficulties regarding shifts and only occurred in two organisations. All participants were given a break between SAMAE and the video to reduce the potential effects of fatigue. It is possible that fatigue could have affected the results, but alternatives could not be made without disrupting work or compromising experimental results. This reflects some of the practical constraints of research in real life environments. However, the limitations inherent in research in real world settings should be weighed against the importance of research in these settings (Koelega, 1996).

Standardised instructions were provided through audio-visual recordings. These explained and demonstrated the TBs and the method of responding, and provided short practice sessions. The core instructions, demonstrations and practice clips were the same for the three treatment groups. In addition, instructions, examples and practice items pertaining to RIGOs and SIGOs were included for those respective groups. The recorded instructions were in English. However, as the first language of some participants was not English, research assistants were available to take questions and explain key aspects of the instructions in the participants' first language where possible. This was necessary to maximise understanding of the instructions.

The practice session was aimed at reducing any initial learning effect, as well as informing participants about the task. Performance in experiments using vigilance tasks tends to increase initially due to practice effects and learning and then reaches a plateau, which is the performance baseline (Pigeau et al., 1995). Instructions and practice lasted thirteen minutes. A lengthier practice session was not considered appropriate as it could have increased fatigue. However, the researcher ensured that all participants were able to respond correctly before beginning the trial. Participants were not given a break during the video as one of the aims of the research was to evaluate vigilance performance over time. Rest pauses could have assisted in reducing fatigue and replenishing attention resources. For ethical and practical reasons, participants were told that the video would last ninety minutes.

The delivery of the video material to participants was standardised at an equipment and software level. Videos were played with VLC media player (http://www.vlcmediaplayer.org/). Laptop computers with dual core processors were used, with simultaneous output to a 15-inch LCD monitor and a CRT monitor. Each participant individually observed their own LCD screen, and the researcher or trained research assistant observed the same display on a CRT monitor. The purpose of using separate monitors for participants and research assistants was to minimise the potential for research assistants to influence the observation process through postural or other cues, recording of responses, or very close physical proximity. The main task of the research assistants was to record the participants' responses and time of response in writing. A time counter appeared on both monitors and was visible to the participants and research assistants. However, the counters were small and were placed in the lower right hand corner of the monitor so as to prevent distraction from the video. Larger time counters were seen as potentially affecting expectancies and the self-regulation of attention resources based on time remaining. In group administration contexts, participants started viewing the video at different times to prevent the chances of mimicking other people's responses.

All groups were required to name the TB as soon as they saw it (pick up, broom flick, kick, and swivel). The research assistant logged the time on the VLC time counter and the type of behaviour on the log sheet (Appendix 4). In addition, the RIGO and SIGO groups were required to respond to IGOs. The RIGO group needed to state 'Shape' when they detected an IGO but were not required to name the shape. Neisser and Becklen (1975) distinguished between perceiving and recognising, with the latter requiring additional processing and attention resources. Therefore naming the shapes is likely to have demanded additional information processing and would have increased distraction from the rest of the display. In addition, it is possible that some participants would have experienced difficulty in naming some of the shapes or objects due to the diverse languages spoken in South Africa. In addition to investigating the importance of semantic distance, RIGOs served the purpose of increasing the frequency of events requiring a response and therefore naming them was not necessary.

Unlike the RIGO group, the SIGO group was required to say 'Picture' as soon as they saw a SIGO and name the TB depicted in the SIGO (e.g., 'Picture kick'). It was necessary to state that a 'picture' was detected so that this could be distinguished from FAs or TBs in the video. Naming the TB depicted in the SIGO indicated that the participant had processed the SIGO to the point of recognising the TB. This was important in demonstrating that sufficient processing had occurred for the participant to recognise the TB so that the SIGO could assist in forming and reinforcing mental models and attention sets for the TBs.

Two sources of qualitative data were gathered with the aim of gaining an understanding of the participants' experiences of the detection process. Firstly, observations of non-verbal behaviours that indicated fatigue and possible attention disengagement were noted. These observations provided qualitative data on the observation process and are elaborated on in Stage 2 where the results are reported. Secondly, a short interview was conducted with participants after watching the video. All participants were asked 'What was it like watching the video?' Participants in the RIGO and SIGO groups were also asked 'What was is like looking for shapes / pictures?' and 'Do you think shapes / pictures could help operators in their jobs? Why/ Why not?' These data are reported with the second study so that comparisons between novices and operators can be made.

6.2.1. Allocation of participants to experimental groups.

The main criterion for allocating participants to experimental groups was their vigilance and observation ability. Therefore assignment to experimental groups was deliberate and not random. This was assessed by means of the Surveillance and Monitoring Assessment Exercise (SAMAE), a computer-based exercise designed specifically for CCTV operators by Leaderware cc (Donald, 1998a). The constructs measured by SAMAE are scanning and dynamic attention, measured through the scanning, dynamic attention exercises taking a total administration time of 90 minutes including instructions and a rest break. These two exercises are considered sufficient in predicting surveillance performance (Donald, 1998b). The scanning exercise 'evaluates a person's ability to rapidly scan a situation and detect a range of subtle to obvious deviations from a defined standard' (Donald, 1998a, p. 1). The dynamic attention exercise evaluates a person's ability to effectively monitor and deal with a situation where attention must be allocated to independent and changing sources of information' (Donald, 1998, p. 1). The computer automatically logs responses. SAMAE items consist of generic shapes and objects that are not based on operational contexts or CCTV footage. Therefore SAMAE does not require SA in order to score well. Rather, it examines a person's underlying ability to observe and visually analyse images and to do this in a dynamic context (Donald, 1998a). Developed in South Africa as a response to security needs around precious minerals, SAMAE is used in a number of countries (United Kingdom, Canada, Australia, Hong Kong, India, Poland, Namibia, Botswana, and South Africa) and sectors including aviation, town centres, casinos, banking, police, customs, and manufacturing and commercial enterprises. An adapted version has been used and tested for the selection of X-ray screeners (ScreenX) and is authorised for usage following a UK Government sponsored validation study (Department for Transport, UK, 2003) that found correlations of .59 with measured threat image projection performance (TIP) on the job.

SAMAE demonstrates acceptable reliability with Kuder Richardson reliability coefficients of .83 and .86 for the scanning and dynamic attention exercises respectively (Donald, 1998b). SAMAE scores are strongly related to various indicators of performance on real world surveillance tasks, such as monthly recorded detection rates, performance appraisals, overall management evaluation scores and non-verbal behaviour recognition training assessments (Donald, 1998a). Therefore predictive validity is high. ' Those scoring in the top 50 percent consistently detected more than double the number of confirmed incidents than those scoring below the average' (Donald, 1998a, p. 5). In addition, SAMAE demonstrates differential validity, as it does not correlate with performance on tasks not related to surveillance (Donald, 1998b).

Given SAMAE's validity, the lack of objective detection measurement data at many operations, and the lack of evidence that biographical variables are related to CCTV detection performance, SAMAE was considered the best criterion for creating matched groups. Therefore SAMAE was the main criterion for allocating participants to experimental groups. Prior to explaining how SAMAE scores were used for group allocation, it is necessary to explain the scores generated. Eight SAMAE scores are generated for each participant: scanning accuracy of detection within first four seconds of items, scanning accuracy of detection within the available time of fifteen seconds, reaction times and FAs for scanning and dynamic attention, accuracy of detection within four seconds for dynamic attention, and a ranked score (Donald, 1998a). For the current study, scanning accuracy for fifteen seconds was considered more important than the four second scanning accuracy score because events on CCTV take a number of seconds to detect and report. The four second scanning accuracy is applicable to jobs such as baggage screeners because aviation x-ray screening typically takes on average about five seconds per bag and screeners have to detect threats within this period. The ranked score is based on a combination of standardised scores of detections within fifteen seconds and FA scores for both sub-tests.

A spread of SAMAE scores within each group was obtained by allocating participant scores to one of four quadrants: greater than or equal to the cut-off point for scanning and dynamic attention; less than the cut-off point for scanning and dynamic attention; greater than or equal to the cut-off point for scanning but not dynamic attention; and greater than or equal to the cut-off point for dynamic attention but not scanning. This was an efficient way of allocating participants to groups in the short time available. Cut-off scores for decisions regarding high versus low scores were based on international norms supplied by Leaderware cc. (i.e., 69.05 for scanning and 50 for dynamic attention, and 5 and 9 or fewer FAs for scanning and dynamic attention respectively). Descriptive statistics for SAMAE scores per group and the total sample are presented in Table 7.

Table 7. Summary statistics of SAMAE scores by treatment group for operators only

Group	Variable	Mean	SD	Min	Max
Benchmark (n=14)	Scan acc 4 secs	32.99	18.32	7.14	59.52
	Scan acc 15 secs	68.71	16.09	28.57	88.1
	Scan RT	4.90	1.17	3.22	6.53
	Scan FA	3.64	3.93	0	13
	Dyn Attn acc	46.93	15.76	13.89	68.06
	Dyn Attn RT	2.09	0.35	1.63	2.75
	Dyn Attn FA	5.07	2.97	1	10
	SAMAE 15 sec ranking	44.73	4.72	36.1	52.2
RIGO (n=15)	Scan acc 4 secs	42.54	17.54	14.29	88.1
	Scan acc 15 secs	72.70	9.85	59.52	95.22
	Scan RT	4.48	1.34	1.96	7.29
	Scan FA	6.40	7.13	0	21
	Dyn Attn acc	48.63	17.90	12.5	77.78
	Dyn Attn RT	2.14	0.23	1.66	2.52
	Dyn Attn FA	11.87	13.84	1	51
	SAMAE 15 sec ranking	45.22	5.69	38.9	59.65
SIGO (n=13)	Scan acc 4 secs	38.46	15.88	11.9	64.29
	Scan acc 15 secs	73.63	12.99	50	90.48
	Scan RT	4.71	1.09	3.21	6.78
	Scan FA	5.46	5.01	0	15
	Dyn Attn acc	55.02	14.84	27.78	80.56
	Dyn Attn RT	2.05	0.27	1.7	2.75
	Dyn Attn FA	12.92	12.85	1	43
	SAMAE 15 sec ranking	45.49	6.70	35.2	54.6

A review of the descriptive statistics for SAMAE scores indicates variations between groups in means for some scores. Therefore linear model analyses of variance (ANOVAs) were conducted to determine whether there were significant differences between groups (Table 8). No significant differences were found between the groups on any of the SAMAE scores.

Therefore the treatment groups are considered equivalent in terms of underlying observation ability based on SAMAE scores.

Table 8. ANOVAs for groups on SAMAE scores

Variable	Model df	Error df	F	p
Scanning accuracy 4 secs	2	39	1.1	0.34
Scanning accuracy 15 secs	2	39	0.54	0.58
Scanning reaction time	2	39	0.43	0.65
Scanning FAs	2	39	0.91	0.41
Dynamic attention accuracy 4 secs	2	39	0.92	0.41
Dynamic attention reaction time	2	39	0.32	0.73
Dynamic attention FAs	2	39	2.05	0.14
SAMAE ranking	2	39	0.06	0.94

* p<0.05, ** p<0.01

In addition to SAMAE, an effort was made to match groups on biographical details that may affect performance (organisation, length of experience as a CCTV operator, job grade, training attended, types of surveillance performed, age and gender (Tables 5 and 6).

The use of SAMAE in allocating participants to experimental groups was substantiated by the significant correlations between SAMAE scores and the total TBs detected and FAs reported in the current research (Table 9).

Table 9. Correlations between SAMAE scores and total TBs and FAs (N=42)

	Scanning accuracy 4 secs	Scanning accuracy 15 secs	Scanning reaction time	Scanning FA	Dynamic Attention accuracy	Dynamic Attention reaction time	Dynamic Attention FA	SAMAE 15 sec ranking
Total TBs	0.55**	0.51**	-0.47**	0.018	0.56**	-0.63**	-0.19	0.54**
	<.0001	<.0001	<.0001	0.91	<.0001	<.0001	0.22	0.0002
Total FAs	-0.21	-0.2	0.14565	0.35*	-0.21	0.14	0.29	-0.39*
	0.19	0.21	0.3574	0.02	0.18	0.36	0.06	0.01

*p,0.05, **p,0.01

Statistically significant correlation coefficients for SAMAE scores with total TBs detected range from -0.63 to 0.56. These are in line with those obtained for X-ray screeners on the ScanX version of the Scanning Exercise by the Department for Transport, UK (2003) and a technical report on correlations between SAMAE tests and on-the-job CCTV detection outcomes (Donald, 1998a). The negative correlations for Scanning reaction time and Dynamic Attention reaction time indicate that participants who responded more slowly (and possibly cautiously) obtained higher TB detection rates. False alarms on the Scanning test correlated significantly with false alarms in the current research (r=0.35, p=0.02).

6.2.2. Pilot studies: Development of materials

This section describes the development of the video, the IGOs and their insertion. The video was the resource used to examine the hypotheses. Before the main study could be implemented, the video and IGOs were developed. Unlike traditional vigilance where simple stimuli are used and their detection difficulty levels can be assumed to be similar, significant events on CCTV differ greatly in salience, duration, and detection difficulty. Therefore it was important to ensure that time phases in the final video were equivalent in terms of factors such as visual complexity and detection difficulty levels. The video development process involved two pilot studies.

The video consisting of a mixture of clips with and without TBs was developed through a series of preparatory steps followed by two pilot studies. The first step was to choose a setting for the video. It was decided that all clips should be based on the same setting or context, ensuring that SA pertaining to only one context is required. A diamond processing plant was chosen as it provided significant events with varying detection difficulty levels and a variety of scenes. Scenes typically depicted one to three employees working in the plant at large machines. They performed activities such as sweeping and shovelling spillage, inspecting machines, hosing areas, filling containers, talking to each other, and being X-rayed as part of security screening. All employees in the videos wore the company uniform and relevant safety equipment. A mixture of actual footage and footage with enacted simulations of theft activities was used.

The second step involved deciding what aspects of significant events should be detected and measured. Many significant events take a while to set up and unfold or involve a series of actions (Donald, 2001). In these situations, it is unclear as to which actions and times should be regarded as the significant event. It was decided that visible, specific TBs consisting of a single action and associated with incidents would be identified. Based on a review and analysis of clips from diamond processing plants, the following behaviours were identified as being frequently associated with significant events: picking up with the hand, kicking and dragging with the foot, stepping on an object and swivelling the ankle sideways to grind the object into the underside of the shoe, flicking an object with fingers, flicking an object with tweezers, and flicking an object to another location with a broom while sweeping. These were therefore considered to be possible TBs or significant events for the study. These behaviours were identified in clips in preparation for pilot study 1. Clips where these behaviours occurred without being associated with irregular events (e.g., picking up tools) were edited prior to pilot study 1. All clips involving flicking objects with tweezers were removed due to quality and difficulty issues.

The third step involved rating video clips in terms of their quality (e.g., focus, lighting), recording the number of people in them, and time at which TBs occurred. Lastly, criteria that were seen as allowing for consistent and equivalent measurement over time in the final video were identified. These are discussed in conjunction with the two pilot studies that were conducted.

6.2.2.1. Pilot study 1: Video content.

The first pilot study assisted in selecting video clips for use in the main study and ensuring that the clips were equivalent across the different time phases of the final video, as elaborated upon later. It consisted of two stages, the first focusing on clips with TBs and the second on non-TB clips. No IGOs were used in pilot study 1.

The clips contained a mixture of footage taken in a real diamond processing operation, and footage created by professional actors in a diamond processing plant. As there was insufficient real life footage of sufficiently high quality, the acted footage was used to supplement it. The acted footage was based on real TBs and significant events and was developed by experts in the area. Filming took place in a diamond processing plant, and was realistic in all respects. Professional actors wore the same uniforms as employees and reflected the typical demographic characteristics of employees in terms of ethnic membership, age and gender. They performed tasks comparable to those undertaken by employees, such as inspecting machines, sweeping and shovelling spillage, hosing areas, cleaning and fixing machines, and going through a full body X-ray. The TBs were interspersed amongst the non-TBs. The acted footage was realistic in all respects and cannot be visually distinguished from the real footage.

In preparation for pilot study 1, the researcher observed approximately 100 TB and 85 non-TB video clips and recorded the following information: file name, duration, place (e.g., type of machine or section of plant), activity occurring, behaviours likely to cause FAs and their time of occurrence, quality of video clip (e.g., focus, black and white versus colour, lighting), and number of people in the clip. For TB clips the TB description and time were recorded and the detection difficulty level was rated by the researcher (1 = extremely easy, 5 = extremely difficult).

The number of people present in the video was considered important as this influences detection difficulty levels and the complexity of the search task. When more than one person appears in a clip, participants divide their attention between people (if close enough on display to do so), switch attention between figures, or focus on particular people to the exclusion of others. Visual complexity of the scenes and salience of the TBs were subsumed under the detection difficulty rating.

A second person performed the same analysis independently and results of the two raters were compared. Consensus was reached through discussion regarding which clips were suitable for inclusion in the pilot study. Clips that contained TBs that were highly unlikely to be detected by observers or which were of poor quality were discarded. Approximately 60 TB and 85 non-TB clips were selected for pilot study 1.

6.2.2.1.1. Pilot study 1a: Target behaviour clips.

Before showing the clips, the TBs were explained, different ways of carrying out the behaviours were discussed, and a demonstration clip of each TB was shown. Participants were then shown the clips as separate files rather than as a continuous video stream. This enabled them to have breaks between clips and to replay clips when necessary. They rated the detection difficulty levels (1 = easy, 5 = extremely difficult, 6 = not detected after one or two replays). FAs were also recorded. Results were analysed by averaging the detection difficulty levels across participants for each TB and clip.

Based on pilot study 1a, a number of clips were eliminated from further use. These included clips with TBs that were not detected by at least four participants and which were therefore considered too difficult; poor quality clips, and clips with more than four people in them. In addition, clips containing two or more TBs that were too close together in time, and those with very high numbers of FAs, were edited or eliminated. The final outcome of pilot study 1a was a list of suitable TB clips categorised by TB type, difficulty rating, area of the processing plant, and number of people.

The pilot study highlighted the fact that all significant events should be visible on the video so that participants were not required to deduce what may have occurred. Deductions are likely to create inaccurate measurement and uncertainty as to what had occurred and at what time. Therefore TBs that occurred out of sight (e.g., obscured from the camera by posture, other people or machines) were eliminated although people with high levels of SA would be likely to deduce that they had occurred.

Pilot study 1a was also useful in identifying TBs that could be used effectively in the research and those that were ambiguous. Finger flicks were extremely difficult to detect and were missed by most participants, even after repeated viewing. Dragging objects under feet were frequently confused with natural movements, creating a high FA rate. Therefore these behaviours were eliminated, leaving picking up, kicking, flicking with a broom, and swivelling the ankle and foot. The detection difficulty ratings were averaged and were used in the allocation of TB clips to time phases in the video for pilot study 2. This ensured that each time phase had equivalent TB clips with similar but varied detection difficulty levels.

6.2.2.1.2. Pilot study 1b: Non-target behaviour clips.

The aim of pilot study 1b was to identify video clips containing behaviours that were likely to be incorrectly perceived as TBs and those that yielded high numbers of FAs. This would prevent clips from inadvertently containing target or ambiguous behaviours and inflating the number of FAs in the final study instead of indicating changes in vigilance.

Participants were three faculty members who observed individual clips and reported any actions that could be perceived as TBs. Results of the participants were compared qualitatively rather than statistically due to the small sample size. Where FAs were noted at a similar time by two or three raters, the clip was edited to remove ambiguous behaviours, or eliminated. Where only one rater reported a FA, the other two raters observed the clip again and through discussion, reached consensus on whether the behaviour was likely to create a FA in the final study. Based on this, a decision was reached as to whether the behaviour should be edited out or the clip eliminated. Editing was performed in most instances. Clips that were blurred or had extremely poor lighting were eliminated. Pilot study 1 provided a useful basis for compiling a video for further testing in pilot study 2.

6.2.2.2. Pilot study 2: Video structure and content

Pilot study 2 was aimed at ensuring that suitable video clips were used in the video for the main study, that there would be equivalence across time phases within the final video, and that instructions to participants were appropriate. Response equivalence across time phases is important because this would be compared in the final video. Therefore it was necessary to ensure that the time phases were equivalent regarding detection difficulty levels and FAs in order to prevent results from being biased and artificially creating detection increments or decrements. This is a fundamental assumption underlying comparison between time phases and differs from experiments using simple visual stimuli against static backgrounds where equivalence can be assumed to be relatively easily.

Three videos were prepared for this pilot study. It was envisaged that revised versions would be combined to form the final video. Each video was thirty minutes long in order to decrease the likelihood of a vigilance decrement from occurring. Although detection performance can decrease in less than thirty minutes (Nuechterlein et al., 1984), it was not practical to request that participants be available on more than three occasions.

Twenty-one TB clips were selected based on pilot study 1 (i.e. seven per thirty minute video). Although this meant that more TBs occurred in the video than would occur in a real life diamond processing plant, a trade-off was made between the length of the experiment and the number of TB clips required to test detection rates. Criteria used in allocating clips were that each video should contain: the same number and types of TBs; similar detection difficulty ratings and proportions of FAs; a range of detection difficulty ratings to allow for the potential effects of the IGO intervention to be measured; a mix of clips both with and without TBs; sufficient time between TBs to ensure that responses could be recorded in writing by the researcher; a variety of plant areas; and a varied number of non-TB clips between TBs.

Detection difficulty ratings were similar across videos, with average ratings ranging from 2 to 2.16 on a six-point scale. Non-TB clips were interspersed between TB clips, with consideration to the areas depicted, duration and number of people in clips. To ensure consistency within videos, the videos were further divided into three ten-minute periods. Two TBs were allocated to two time periods within each video and three to one time period.

Observers develop expectancies regarding when and how frequently significant events occur and this influences their responses (MacMillan & Creelman, 1991). This is particularly relevant to situations where significant events occur regularly or predictably (See et al., 1995). To prevent this from occurring, the number of non-TB clips between TB clips was varied and ranged from zero (where two TBs occurred in the same clip) to six. In addition, TBs occurred at different times within each ten minute time period. The intention was to reduce the likelihood of expectancies regarding the occurrence of significant events from developing.

Five students (aged 23 to 28 years, one male and four females) voluntarily participated in pilot study 2. Due to time constraints, one participant only observed two videos. Videos were administered individually and on different occasions to assist participants in maintaining their concentration. Unlike pilot study 1, clips were merged into continuous video stream which could not be stopped or replayed. No IGOs were used in the second pilot study. To reduce the effects of learning and the vigilance decrement across videos, the videos were administered in a different order for each participant. Instructions were videotaped to ensure standardisation. This included videotaped demonstrations of the TBs and a short practice session.

Participants verbally stated which TB they detected and rated their degree of certainty. The rating was based on fuzzy signal detection theory (Masalonis & Parasuraman, 2003) which allows for differences in detection certainty. A 3-point rating scale was used. The researcher wrote down the participants' verbal responses (TB type, certainty rating), time of occurrence and result (hit, miss, FA, correct rejection). These were recorded on a time line containing a list of the clips, the TBs and the time at which they were supposed to be observed. The time line was considered the 'ground truth' (Andrew et al., 2003) regarding the occurrence of TBs. It also allowed for a moment by moment recording of responses.

Results were analysed in terms of frequencies of response type by ten minute period for each video. Recommendations are discussed in conjunction with the results. Hits consist of reports of the correct TB from the actual time of occurrence up to four seconds later. This time range was found to be necessary in view of participants' speed of response and the time required for verbal responses. A four second interval was chosen based on scrutiny of the results which indicated that virtually all correct responses occurred within this time frame as reaction time was not considered critical for this task. This is therefore seen as an appropriate length of time. Different hit rates were obtained for the three videos (A=77%, B=100% and C=54%), indicating different detection difficulty levels and the need for redistribution of TBs in the final video. However, different detection rates were expected for operators with experience at CCTV surveillance compared with the students who participated in the pilot study.

Clips where all participants detected the TB were considered too easy and were eliminated for the final video. (Similarly, clips that were considered too difficult had been eliminated during pilot study 1a). Conversely, TB clips with low detection rates needed to be spread more evenly between the time phases for the main study. No TBs were missed by all participants, suggesting that none of them was too difficult to detect.

Results indicated that certain clips needed to be redistributed between time periods for the main study. In addition, clips with a large number of FAs were re-observed to ascertain the basis for them. This analysis indicated that ambiguous movements, misinterpretation of actions, and misunderstanding of what constitutes TBs versus natural movements formed the basis of most FAs. Remedial actions included editing out ambiguous movements, altering instructions, and excluding certain clips. A few clips were left unchanged as the certainty ratings were low and there was little visual basis for the FAs. The aim of changing instructions was to clarify what constitutes certain TBs versus natural movements.

The structure of the videos and intervals between TBs seemed appropriate and was seen as appropriate for the final video. Given the fact that the main study would include IGOs, increasing the number of responses required, it was decided to reduce the number of TBs to six per thirty minutes. This remains more than would typically be encountered in the work place, but provides sufficient opportunity to demonstrate the potential effects of IGOs.

The use of videotaped instructions was useful in standardising instructions. The instructions were generally clear and no participants asked questions. However, the instructions were improved for the main study by clarifying what constitutes TBs.

Pilot study 2 was an essential step in creating an appropriate continuous video with appropriate numbers of TBs for the final study. It assisted in identifying clips with TBs that were too easy and those where a high number of FAs were consistently reported. It provided pointers as to how clips both with and without TBs should be divided between time phases in the final video. Lastly, it assisted with improving instructions in order to assist detection and reduce FAs. The results were used in compiling the video for the main study.

6.2.2.3. IGO development and insertion.

The two treatment groups received different sets of IGOs. All IGOs can be regarded as novel because no participants had previously been exposed to them. The key difference between the IGOs was the degree of semantic distance they depicted to TBs. RIGOs had no relationship to significant events or the video's context. They consisted of simple shapes and objects, such as a circle, fish, leaf or ball (figure 1). These objects were chosen because they are similar to the types of images used by researchers in the area (e.g., Andrew et al., 2003; Neil et al., 2007) and therefore provide a basis for comparison with previous studies.

Figure 1. Examples of RIGOs

SIGOs had small semantic distances to TBs and represented both TBs and the postures associated with them. All SIGOs represented one of the TBs used in the current research (pick up, kick, broom flick, and ankle swivel). As these behaviours can be performed in a number of ways that are associated with various postures (e.g., bending over, crouching), it was important to depict the posture as well as the TB. Video clips with TBs that had been selected for the video were scrutinised to identify the relevant postures.

Models were hired and photographed performing the TBs in the relevant postures. These photographs formed the basis of the SIGOs (figure 2).

Figure 2. Examples of SIGOs (Clockwise from top left, the TBs represented are a pick up, swivel, kick, pick up, and broom flick).

Models in the photographs wore the same clothing (i.e., blue overalls, hard hats, safety boots, gloves) and used the same tools (e.g., hammer, broom, shovel) as people in the video, in order to create a context similar to the videos and similar visual salience levels as people in the video. Models were photographed against a green chroma key screen to allow for the elimination of the photographs' backgrounds and insertion into the video. Eighteen photographs were selected to be SIGOs from a total of 200 photographs. These were selected by the researcher, an expert in CCTV surveillance, and a research assistant. The three parties made individual selections of the photographs they perceived to portray the TBs best. Where there was disagreement, discussion took place until consensus was reached. A pilot study aimed at establishing how accurately the SIGOs portrayed the designated TBs was not carried out because the above process was considered thorough and the SIGOs were by nature concrete. Based on researchers such as McDougall et al. (2001), concrete images are more easily comprehended than abstract ones. In addition to small semantic distances, SIGOs were concrete, complex and realistic.

A total of twelve IGOs were inserted per treatment group, with four per thirty-minute phase. These were inserted at varied intervals within the thirty-minute phases to prevent a response set from forming. All IGOs were static but appeared abruptly for four seconds and then disappeared abruptly. RIGOs and SIGOs were inserted at the same times and in the same locations. Locations were not necessarily those where people would typically appear and did not cue the location of the next TB. IGOs were inserted at times and in locations that did not overlap with TBs. RIGOs were smaller than SIGOs because RIGOs tended to dominate the display and increase inconspicuity when they were as large as SIGOs. However, SIGOs needed to be sufficiently large for participants to recognise the TB they represented. In addition, SIGOs that were a great deal smaller than people in the video were highly conspicuous. Therefore SIGOs were similar in size to people in the video and RIGOs were approximately ten percent of the screen height. This is consistent with Andrew et al., (2003) and Neil et al. (2007) who conducted studies on inserted images in CCTV.

Only six SIGOs were similar to the TB that followed it in terms of type and posture of the TB. These were spread evenly between the three time phases. However, this was not intended as a cueing effect and cannot be compared with other vigilance or visual search research because the time that lapsed between the IGOs and the relevant TB was much longer than used in other research. Studies using cues have typically used extremely short times between the cue and the event. These are typically in the order of milliseconds (Richard, Wright, & Ward, 2003) rather than minutes as in the current research. In the current study, the time varied between IGOs and TBs, ranging from 1m12s to 7m3s with a mean of 2m54s. The lengthy interval was used because the purpose of IGOs is not to act as cues that direct the eyes to a particular location, but to influence the underlying attention sets. The ability of stimuli to cue significant events appears limited to extremely short intervals (MacLean et al., 2009) and is therefore unlikely to have applied to the current research as in real settings one would not know when a significant event is taking place in order to cue that event.

These intervals are much longer than those typically used in vigilance research, but are shorter than the intervals between significant events in real life CCV surveillance settings. In operational CCTV surveillance contexts, the frequency of significant events ranges considerably (Donald, 2001).

6.2.2.4. Final video structure

The final video was ninety minutes long. This was deemed long enough to allow for a potential vigilance decrement to set in, and to have sufficient significant events and IGOs while including periods in between with no significant events. Previous research has found that a vigilance decrement typically sets in by 20 to 35 minutes (Sawin & Scerbo, 1995) but under certain conditions the decrement has been known to occur in as little as five minutes (Nuechterlein et al., 1984). Ninety minutes is considerably longer than many vigilance studies and therefore allows for examination of the vigilance process over an extended period.

For the purpose of analysis, the video contained three time phases of thirty minutes each. The same number of TBs and IGOs was contained in each thirty minute phase. To ensure a relatively comparable spread of TBs and IGOs within each phase, the phases were further divided into ten-minute periods, each with the same number of TBs. IGOs were inserted fairly evenly throughout the video, but not so regularly that participants were able to predict when they would occur. Care was taken to ensure that IGOs did not overlap with TBs. IGOs were inserted at the same times and places for the RIGO and SIGO groups. The sequence and times of TBs and IGOs are presented in Appendix 5. Pinnacle Studio 9 Plus video editing software was used to insert IGOs and edit video material. TBs took up to two seconds to unfold and were therefore displayed for a maximum of two seconds. The order of the clips and time phases remained the same for all participants.

6.3. Measurement

Hits, misses and FAs (FAs) formed the units of measurement. These measures are based on SDT (MacMillan & Creelman, 1991). The times at which responses occurred were recorded. However, their purpose was to ensure that responses were correctly categorised as hits, misses or FAs rather than to measure reaction times. Although reaction times are typically used to identify the vigilance decrement (Buck, 1966), this was not possible in the current study due to the recording methods used. Responses were recorded manually by research assistants. A time lag inevitably occurred between the perception, recognition and verbal reporting of significant events. Therefore measurement of reaction times would have been inaccurate. Discussion with several CCTV surveillance experts indicated that TBs and significant events could take several seconds to be detected. Furthermore, operators varied in their speed of reporting the events. Experts emphasised that detecting the events was more important than small differences in detection speed in the CCTV context, even though quick detection was required. Therefore reaction times were not considered important to measure. A trade-off was made between ensuring that the correct visual stimulus was recorded and measuring reaction time.

Various alternative recording methods were considered, including an automated logging system by clicking on a mouse, participants typing or writing the time and nature of their response, and videotaping of responses with participants verbalising their responses and pointing to the display. However, two considerations were paramount. The first was that the recording process should not interfere with the observation process or potential vigilance decrement. If participants had recorded their own responses, in whatever form, the nature of the observation and vigilance task would have changed. Additional tasks and information processing requirements would have been introduced. Based on the job characteristics model (Hackman & Oldham, 1976), greater task variety would also have been introduced with potential influences on motivation and alertness.

The second consideration was that the accuracy of data needed to be ensured. As the displays are dynamic and at times show more than one person, it was important to ensure that responses were recorded correctly. As reports of hits, misses and FAs could occur close together in time, the automation of logs or recordings by participants was likely to create difficulties in terms of interpreting the data. Therefore manual recording by individual research assistants was used as this allowed for clarification of responses when required. However, this had the disadvantage that reaction times could not be measured accurately. This is offset by the greater importance of accurate detection compared with reaction times in real world CCTV surveillance.

The video was designed so that detection accuracy over thirty- and ten-minute periods could be analysed. Some laboratory studies of vigilance have used much shorter intervals (e.g., Neuchterlein et al., 1983). However, shorter periods would have meant that more TBs would have had to be inserted, increasing the frequency of significant events. This would have changed the nature of the task by reducing monotony and would have reduced the similarity between surveillance tasks performed by operators in the real world and this research.

Hits, misses and FAs were logged by the researcher and assistants for IGOs and TBs. The type of TB, RIGO or SIGO that was reported was recorded, as well as the time at which this occurred. Data was converted to a format where the different groups' performance of the groups could be compared over time. This was based on a time line depicting the occurrence of TBs, IGOs, and allowing for the recording of FAs. This approach has been used since early research on vigilance (e.g., Howell et al., 1966). Written logged detections were placed on the time line and compared against 'ground truth' information to ensure that the response reported was a correct detection rather than a FA. In rare instances, participants reported that a picture had been detected, but were unable to decide what TB was depicted, or reported the incorrect TB. These were recorded as 'misinterpretations' and received a different code from hits and misses.

In addition to separate TB and FA scores, an accuracy score was calculated. Detection performance is based on both the detection of TBs and FAs (MacMillan & Creelman, 1991). Therefore these scores were combined into an accuracy score that included both TBs detected and FAs reported. This was calculated by converting TBs and FAs to ranked normalised scores using Blom's method (Altman, 1991). FAs were then subtracted from TBs to provide an accuracy score.

To evaluate the importance of difficulty level in the detection of TBs, a detection difficulty index was estimated, based on the two pilot studies. The index consisted of a combination of difficulty ratings from pilot study 1 and actual detection rates in pilot study 2. In pilot study 1, six participants individually observed the clips and provided difficulty ratings on a one to six scale (1 = easy, 5 = very difficult and 6 = not detected after one or two replays). Participants were able to replay clips and take breaks when required to reduce a potential vigilance decrement and influence on their ratings. Ratings were averaged for each TB. Clips with an average rating of six were discarded and were not included in the final video. In pilot study 2, five participants watched each thirty-minute phase on a different occasion from the other thirty-minute phases, again attempting to reduce the vigilance decrement. In addition, the order in which participants watched the phases was varied. The percentage of participants who detected the TBs was calculated.

Difficulty level ratings and detection levels from the two pilot studies were compared and combined into a difficulty level index (Table 10). The index provided a gross estimate of the difficulty level and consisted of 1 = easy, 2 = moderate and 3 = difficult.

Table 10. Difficulty index for TBs

	Pilot study 1	**Pilot Study 2**	
TB	**Rating (mean)**	**Detected (%)**	**Index**
1	1.8	80	1
2	2.2	60	2
3	4	40	3
4	4.3	40	3
5	2	40	3
6	2.3	100	1
7	4	40	3
8	3.3	80	1
9	1.5	100	1
10	1.7	N/A	3
11	3	40	3
12	3.3	40	3
13	2.8	100	1
14	1.5	40	3
15	4	40	3
16	1.5	80	1
17	2.5	80	1
18	4	40	3

Note: N/A indicates that TB was not included in pilot study

The difficulty indices were summed for each relevant time period. Each thirty-minute phase included four TBs, and each ten-minute period, two TBs. This allowed for a comparable index for thirty- and ten-minute time periods.

The difficulty level index is an approximation and has its limitations. Firstly, the sample sizes in the pilot studies are limited. Secondly, the 3-point scale does not provide a detailed analysis of difficulty level, but merely a gross differentiation between 'easy,' 'moderate' and 'difficult.' However, given the small sample sizes in the pilot studies, a more fine-tuned analysis was not appropriate. Thirdly, only one TB was allocated a 'moderate' difficulty level, effectively limiting analyses to 'easy' versus 'difficult' TBs. Lastly, the difficulty levels of two TBs were estimated by the researcher. TB 10 had not been included in pilot study 2 as it was not considered suitable. However, an edited version was included in the final video. Therefore the researcher estimated its difficulty index based on pilot study 1 and the types of TBs that participants detected and missed in pilot study 2. TB number 14 showed a discrepancy between the two pilot studies and the researcher therefore estimated the rating based on knowledge of this TB in relation to others. Despite these limitations, reasonable care was taken to ensure that the difficulty index has adequate validity.

6.4. Analyses

The hypothesis regarding overall detection performance (based on TBs, FAs, accuracy, and IGOs) on the designated task was examined using percentages, means and standard deviations. The remaining hypotheses were examined by means of repeated measures analyses of covariance (ANCOVAs) and repeated measures two- and three-way analyses of variance (ANOVAs) using mixed procedures and between subjects factors. ANCOVAs were only used to evaluate the contribution of the TB difficulty index (explained in section 6.3). Where the difficulty level was significant, an adjusted TB score was used that took difficulty level into account and two-way ANOVAs were used. ANOVAs were calculated for FAs and IGOs. Unless otherwise stated, the results refer to repeated measures two or three-way ANOVAs.

Repeated measures ANOVAs and ANCOVAs allow for the comparison of differences in means over several measures for the same participants and between several groups (Rosenthal & Rosnow, 2008). The variable that was repeated in the current research was measurement over the thirty-minute phases (for TBs, FAs, accuracy and IGOs). These constituted the within-subjects measures. The between-subjects measure consisted of the experimental groups and for Stage 2, surveillance background was also included. With ANOVAs, comparisons of means for different groups is achieved by subdividing the total sum of squares between components to which the observed variance can be attributed

(Clarke & Cooke, 2004). ANOVAs perform multiple comparisons of means by comparing the residual variability that remains once means have been removed. ANOVAs use Fisher's F-distribution to test whether the means of groups are equal. This produces an F-value which is the ratio of estimated variance in random samples from normal distributions. The F-value and its associated degrees of freedom are associated with p-values which determine significance. For the current study, 0.05 was the chosen level of significance or Alpha value, although the higher level of .01 was noted where applicable. The mixed procedure uses a number of mixed linear models to fit data, is flexible and is suited to repeated measures analyses (SAS Enterprise Guide, n.d.). Where necessary, (i.e. where significant differences were found in the relevant ANOVA/ANCOVA), Tukey's Honestly Significant Differences (HSD) tests were conducted to identify where significant differences in means occurred (Rosenthal & Rosnow, 1985).

The responses from the interviews for both studies were analysed by means of thematic content analysis and frequency counts (Smith, 2000). The combined use of qualitative thematic content analysis and quantitative frequencies provided indications of the types of responses given and how commonly these occurred. In addition to interviews, observation of participants by the research assistants provided qualitative data. Research assistants logged the time and nature of non-verbal behaviours (e.g., looking away from the display, closing the eyes, fidgeting and other displacement behaviour, and a 'glazed' expression on participant's faces). These actions were interpreted as indicating attentional disengagement from the task provided that participants did not report any events (hits or FAs) during this time. The researcher then identified 'disengagement periods' by perusing the logs and identifying periods where notes about non-verbal behaviours corresponded with a lack of responses for several consecutive minutes. Periods with no responses that were not accompanied by notes regarding non-verbal behaviours were not considered as disengagement periods, because the participant may have been concentrating but not observing anything to report. The start and end times of the relevant clips were used as a rough indication of the length of the disengagement periods. Disengagement periods were difficult to identify for participants who reported very few FAs, and are therefore probably underestimated for these individuals.

Based on Cheyne et al.'s (2009) model of disengagement, only the third level of disengagement was measured. These periods of disengagement were accompanied by signs of motor restlessness and fatigue. Alternative procedures would have been necessary to evaluate disengagement that was not accompanied by external signs. The use of deeper levels of disengagement is therefore likely to have resulted in conservative estimates of disengagement. It is emphasised that the identification and calculation of disengagement periods is a crude measure and was not supported by neurophysiological measures that could have provided more accurate measurement. The number and length of disengagement periods are mere estimates and all references to them should be treated with caution. However, they were seen as providing a tentative exploration of attention engagement during the vigilance task and are therefore retained in the results.

The methods used for identifying periods of disengagement could be criticised on the grounds that it appears to mirror the measurement of misses, reducing its validity. However, care was taken to ensure that physical signs of restlessness, fatigue or disengagement accompanied the disengagement periods. Further, disengagement is a process which is different from hits and misses which are task outcomes.

Secondary analyses of the biographical variables were aimed at identifying potential explanations of the results. The relationships between biographic factors, TB detection and FAs were analysed by means of ANOVAs and t-tests. Where variances were unequal, the Satterthwaite method was applied, otherwise the pooled method was used. All statistical analyses were conducted with SAS Enterprise Guide version 4.2 (2008) and SAS version 9.2 (2008).

6.5. Participant involvement and ethical issues

The research was approved by the appropriate university ethics committee (Appendix 6). Written permission was obtained from the organisations regarding access to CCTV staff. Participants were given a letter informing them of the purpose of the research, the research process, requirements including time involved, and participants' rights in lay terms (Appendix 2). The invitation was issued to potential participants independently of managers or supervisors. Participations signed consent forms regarding participation in the research, the use of SAMAE and the provision of SAMAE scores to their organisations (Appendix 3).

Participation was voluntary. Individual results are not reported in the thesis and only general trends are presented. Similarly, no organisational identities are disclosed. The research required four to five hours per participant. As this is an extensive period of time, breaks were provided. The large demand on time was possibly the greatest drawback in obtaining a sample. Organisations re-scheduled shifts and off-days in order to accommodate the research. The trials were conducted at the operations in quiet venues.

Research assistants involved in collecting data were fully trained and supervised. The researcher was available at all times to assist with problems and answer queries. In addition, the participants were fully briefed and trained on the procedures and responses required, and opportunities to practice the responses, ask questions and obtain assistance were given. Time for these activities is included in the times mentioned above.

In the event of participants requesting feedback, summaries of the overall research results were provided to them via email, fax or mail. Feedback on individual performance on the video was not provided to any parties as the research aims to identify suitable characteristics of IGOs, rather than individual performance on visual detection tasks. Participating organisations were provided with an executive summary and a full report of the results.

In summary, the research consisted of two stages, with the same research design, procedures and video being used at both stages. Stage 1 included 42 operators and Stage 2 extended the sample to include an additional 31 novices, yielding a total sample of 73. In Stage 1, participants were treated as a single, global group, whereas in Stage 2 the sample was divided into sub-samples of novices, generalists and specialists, allowing for the identification of different vigilance dynamics between participants with different types and degrees of exposure to CCTV surveillance. The next chapter presents the results for the two stages.

CHAPTER 7: RESULTS

This chapter provides the results of the research. Stage 1 refers to the initial sample operators, treated as a single sample. The hypothesis regarding the level of detection performance is examined in terms of target behaviour (TB) detection rates, false alarms (FAs) and accuracy. Findings regarding potential vigilance decrements are presented next, also in terms of TB detection rates, FAs and accuracy. Following this, the effects of the IGO intervention are examined. These are followed by secondary analyses of biographical data aimed at identifying variables that explain the findings. The emergence of surveillance background as a factor that is likely to explain variations in vigilance dynamics is explained at this point. Results are then summarised and the rationale is provided for extending the sample to include novices and for examining the data in terms of different surveillance backgrounds.

Stage 2 presents the results for the initial sample of operators (N=42) as well as novices (N=31), giving a global sample of 73. In the introduction to Stage 2, amended hypotheses are presented. These incorporate surveillance background, but remain focused on the level of detection performance, potential vigilance decrements and the effectiveness of the IGO intervention. The results are then presented, following a similar order to that adopted in Stage 1. For each hypothesis, results are first presented for the global sample of 73, then are examined in terms of the three strata, with operators categorised into the sub-samples of generalists and specialists, and students forming the novice sub-sample. These groupings allow surveillance background to be taken into account. Surveillance background includes the type of CCTV surveillance tasks to which participants have been exposed and is associated with length of experience working in CCTV surveillance and training in CCTV systems and surveillance. While novices have no CCTV surveillance experience or training, operators are categorised on this basis as explained at the end of the results for Stage 1.

For both stages, results are presented for each hypothesis in turn. Descriptive statistics are provided at relevant points within each hypothesis.

The reader is reminded of the list of non-standard abbreviations in section 1.4.

7.1. Stage 1: Operators

This section provides the results only for participants who worked as operators (N=42).

7.1.1. Examination of hypotheses.

Hypothesis 1: The rate and accuracy of detection of realistic target behaviours (TBs) in a complex CCTV surveillance work environment is lower than for known and visually salient significant events in controlled CCTV settings.

Table 11 presents the summary statistics for the sample as a whole.

Table 11. Summary statistics on dependent variables for operators (N=42)

Variable	Mean	SD	Min	Max
TBs Phase A	3.50	1.38	1	6
TBs Phase B	2.86	1.18	1	5
TBs Phase C	3.60	1.52	1	6
Easy TBs	4.71	1.69	1	7
Difficult TBs	4.57	1.86	1	9
Total TBs	9.95	3.23	4	17
IGOs Phase A	6.89	1.64	2	8
IGOs Phase B	6.93	1.25	4	8
IGOs Phase C	7.11	1.17	4	8
Total IGOs	20.93	2.89	12	24
FAs Phase A	6.71	5.86	0	26
FAs Phase B	4.43	4.35	0	18
FA Phase C	4.69	3.65	0	16
Total FAs	15.76	11.92	0	46

The summary statistics are used to draw conclusions about the level of detection performance achieved by the sample on the designated ninety-minute vigilance task.

The mean number of TBs detected by operators was 9.95. This is slightly over half (55%) the TBs in the video. The minimum was four (22.22%), and the maximum was 17 (94.44%). The standard deviation of 3.23 indicates that most operators detected between 6.72 (37.33%) and 13.18 (73.22%) TBs, indicating high variance regarding detection rates. The mean for TBs rated as easy was 4.71 and for difficult TBs was 4.57. As there were seven easy and ten difficult TBs, this shows that a greater proportion of easy (67.29%) than difficult (45.7%) TBs was detected. Additional scrutiny of the data revealed that only five (11.9%) operators detected fourteen (77%) or more TBs. No operators detected all TBs. Six operators

(14.29%) detected six (33.33%) or fewer TBs. This demonstrates that a substantial proportion of TBs was missed.

The mean number of FAs reported for the whole video was 15.76, with a range from zero to 46 and a standard deviation of 11.92. This shows considerable variance. Relatively few operators reported 24 or more FAs. However, the FA rate for operators as a whole on the video was high. Most participants reported the majority of IGOs (mean score=20.93, SD=2.89), providing an opportunity for the intervention to have an effect. The percentages of RIGOs (85,54%) and SIGOs (89.08%) detected were high. This is not surprising given the visual conspicuity of the IGOs.

Overall, the low TB detection rates indicate that a large proportion of TBs was missed. Detection accuracy was further reduced by the high level of FAs. The detection rate and accuracy for the CCTV surveillance task used in this research was therefore lower than those reported in studies where the significant events were known and visually salient. This provides support for the first hypothesis.

Hypothesis 2: There is a vigilance decrement over time on the designated ninety-minute task.

The existence of a vigilance decrement was tested for the sample as a whole and independently for each experimental group. Because the experimental groups' performance could differ due to potential effects of the intervention, the vigilance decrement was most likely to be found in the benchmark group. The vigilance decrement was examined in terms of TBs detected, FAs, accuracy, and detection of IGOs across experimental conditions. TB means and standard deviations for each time phase and experimental group are provided in Table 12.

Table 12. Descriptive statistics for TBs by phase and experimental group (N=42)

		Phase A		**Phase B**		**Phase C**	
Group	**N**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Benchmark	14	3.36	1.45	2.86	1.41	3.50	1.87
RIGO	15	3.60	1.45	2.93	1.10	3.40	1.30
SIGO	13	3.54	1.33	2.77	1.09	3.92	1.38
All operators	42	3.50	1.38	2.86	1.18	3.60	1.52

Repeated measures ANOVAs for thirty-minute periods found that phase was not significant for TBs ($F(2, 74.6)=1.32$, $p=0.27$), indicating that the means for TBs detected were similar over time.

The experimental group was not significant for TBs ($F(2, 38.7)=0.08$, $p=0.92$) and neither were the interactions of group with phase ($F(4, 75.6)=0.47$, $p=0.76$). Therefore the benchmark group did not achieve a different pattern of TBs over time from the treatment groups. These findings are presented graphically in figure 3.

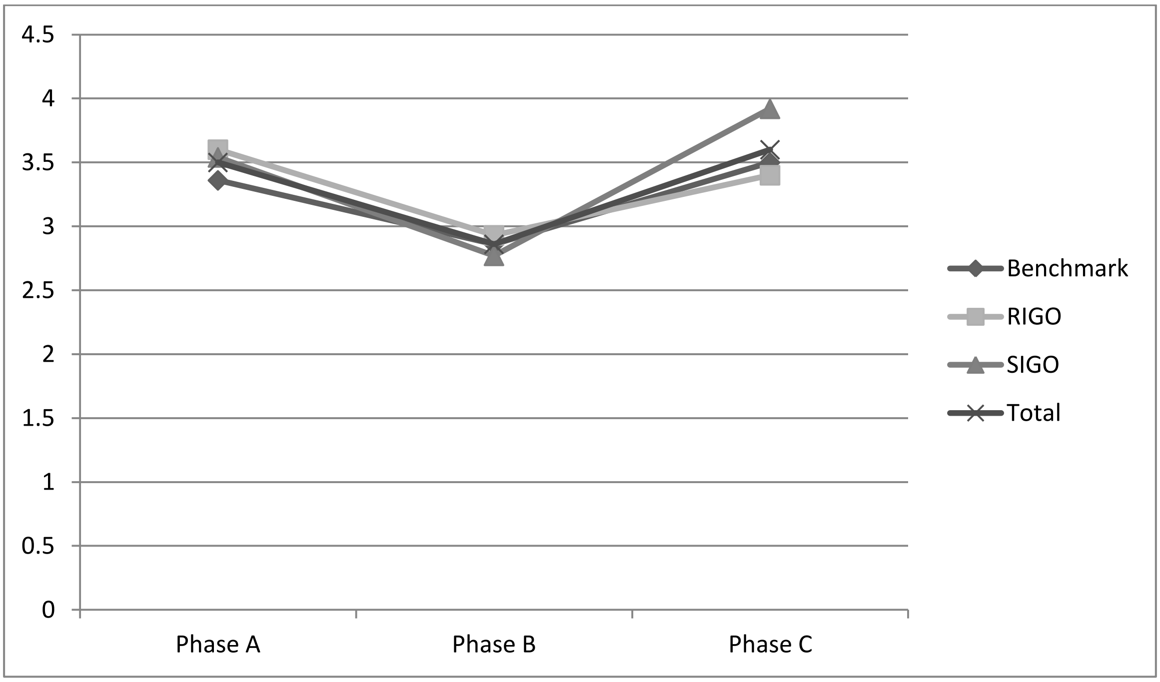

Figure 3. TBs by experimental group and phase (N=42)

The above figure illustrates variations in TB detection over time, with a decrease followed by an increase. However, these variations were not significant. Results for the experimental groups are similar. Although the SIGO group's TB detection level increased more than the other groups' in Phase C, this was not significant and therefore did not show an effect.

The means and standard deviations for FAs per phase and group are presented in Table 13.

Table 13. Summary statistics of FAs by phase and experimental group (N=42)

		Phase A		Phase B		Phase C	
Group	**N**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Benchmark	14	5.79	4.12	4.36	2.62	4.50	2.24
RIGO	15	6.27	4.95	4.20	4.06	5.00	4.09
SIGO	13	8.23	8.17	4.77	6.15	4.54	4.52
All participants	42	6.71	5.86	4.43	4.35	4.69	3.65

FAs were significantly different for different phases (F(2, 76.4)=7.69, p=0.0009), but not for experimental groups (F(2, 39.5)=0.19, p=0.83), or the interaction between phase and experimental group (F(4, 76.9)=0.68, p=0.61). Comparisons of mean FAs between phases are presented in Table 14.

Table 14. Differences of least squares means for FAs between phases (N=42)

Effect	Phase	Phase	Estimate	SE	DF	t Value	Pr > |t|	Lower	Upper
Phase	A	B	2.32	0.59	75.3	3.92**	0.0002	1.14	3.5
Phase	A	C	2.08	0.77	97.3	2.7**	0.01	0.55	3.61
Phase	B	C	-0.24	0.59	75.3	-0.4	0.69	-1.42	0.94

*p<0.05, **p<0.01

The means for FAs for the first phase were significantly different from the second (t=3.92, p=0.0002) and third (t=2.70, p=0.01) phases. However, the mean FAs were similar for the second and third phases (t=-0.40, p=0.69). Examination of the estimated means indicates that FAs decreased from the first (6.76) to second (4.44) phase and were similar for the second and third (4.68) phases. The trends for FAs between phases for experimental groups are illustrated in figure 4.

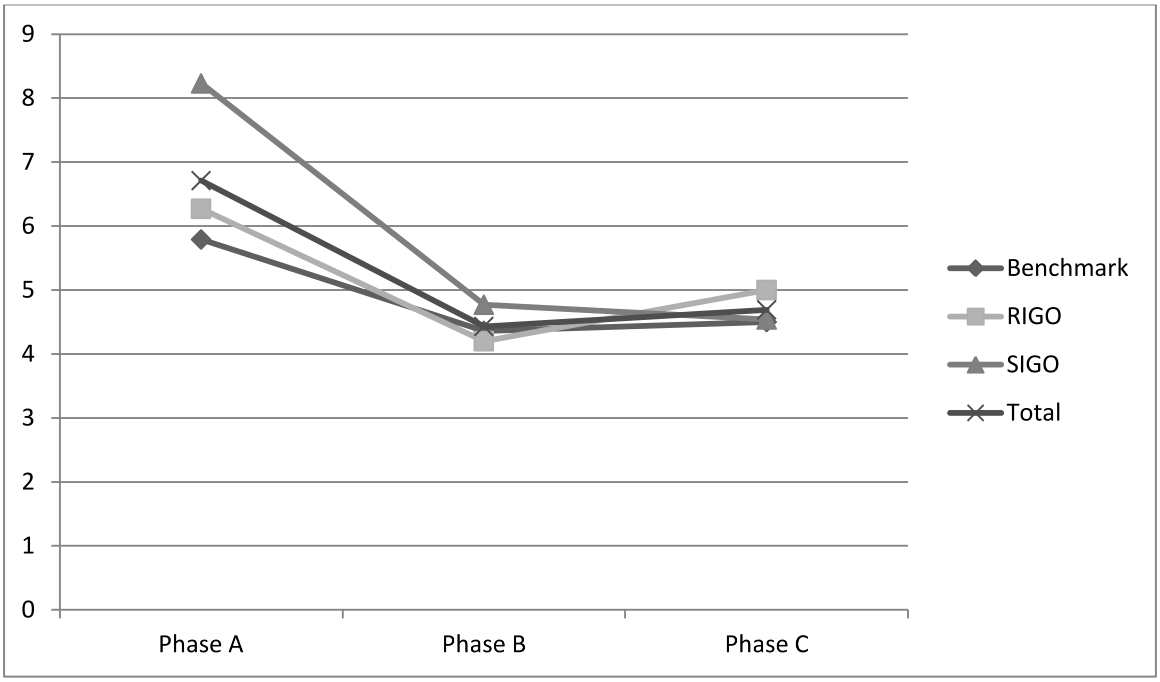

Figure 4. FAs by experimental group and phase (N=42)

FAs decreased significantly for all groups from Phase A to B and then leveled off. From the above figure it appears that FAs were higher for the SIGO than other experimental groups at Phase A, but these differences were not significant.

When FAs and TBs were combined into an accuracy score, phase was non-significant ($F(2, 77.1)=0.91$, $p=0.41$).

The detection of IGOs over time was also examined. The means and standard deviations for IGOs are presented in Table 15.

Table 15. Summary statistics of IGOs per phase (N=42)

	Phase A		**Phase B**		**Phase C**		**Total IGOs**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
RIGO	6.8	1.82	6.93	1.28	6.8	1.26	20.53	3.34
SIGO	7	1.47	6.92	1.26	7.46	0.97	21.38	2.33

Phase was not significant for IGOs ($F(2, 50.4)=0.25$, $p=0.78$), indicating consistency in IGO detection over time.

FAs are the only variable that showed a decrease over time. Taken together, however, the above results tend not to support the hypothesis that a vigilance decrement occurred across the entire 90-minute period. An alternate interpretation is that there was a vigilance decrement (from Phase A to Phase B) followed by a motivational mobilisation of attention in Phase C. Since participants were able to see the time on their monitors, this seems like a plausible explanation.

Hypothesis 3: IGOs enhance detection performance over time, with SIGOs being more effective than RIGOs.

Before analysing results for the hypothesis that IGOs enhance detection performance, it was important to ensure that enough IGOs were detected to create the opportunity for the intervention to have an effect on TB detection. Based on the codes allocated to IGOs, the maximum possible total score was 24. Most IGOs were detected by the majority of participants (overall mean=21.38, SD=2.33) (Table 15), providing the basis for IGOs to make an impact. Further, the mean number of RIGOs and SIGOs detected was similar (t=-0.82, p=0.42).

Summary statistics for TBs per experimental group were provided in Table 12. The treatment groups were not significantly different from the benchmark group on the detection of TBs (F(2, 38.7)=0.08, p=0.92). Comparisons of the benchmark groups' means for TBs with those of the RIGO (t=-0.18, p=0.86) and SIGO (t=-0.41, p=68) groups separately confirmed that the two IGO groups were similar in TB detection to the benchmark group. In addition, the RIGO and SIGO groups' means were similar to each other (t=-0.24, p=0.81).

Moving to FAs, summary statistics for total FAs were provided in Table 13. It is noted that the standard deviations are extremely high for FAs, indicating large individual variability. The treatment groups were not significantly different from the benchmark group (F(2, 39.5)=0.19, p=0.83), nor were the interactions between group and phase significantly different from each other (F(4, 76.9)=0.68, p=0.61).

There were no significant differences between the treatment and benchmark groups on accuracy (F(2, 40.4)=0.10, p=0.91), indicating that IGOs had no effect when TBs and FAs were considered together.

The results consistently fail to support the hypothesis that IGOs enhanced the detection of TBs.

7.1.2. Secondary analyses.

In view of the lack of significant results for hypotheses regarding the vigilance decrement and the intervention, biographic variables were analysed with the aim of identifying factors that could explain the results. Data for the experimental groups were collapsed as no significant differences had been found between them. For TBs detected, significant differences were found for age ($r=-0.45$, $p=0.003$) and race ($F(2, 39)=4.69$, $p=0.02$). Younger participants tended to detect more TBs than older participants. White participants detected significantly more TBs than Black participants, but similar numbers to Coloured participants. No significant differences were found on TB detection rates for organisation, CCTV experience, surveillance type (live, review, or a combination), training type, education level, or gender. None of the variables was significant for FAs or accuracy.

The demographic variables provided little in the way of explanation of the variance in detection performance. The researcher therefore speculated that more subtle differences in the types of surveillance tasks conducted by participants may exist. The type of surveillance conducted by the six organisations was discussed in conjunction with personnel from the organisations and a specialist in the area. It emerged that three organisations performed a general, largely alarm- and response-based surveillance function, while the remaining three organisations performed more proactive surveillance that relied more heavily on operators' detection skills and SA. These are referred to as 'general' and 'specialised' surveillance respectively for the purpose of this research and participants as 'generalists' and 'specialists.'

General surveillance consisted of relatively strong reliance on auditory alarms to indicate significant events rather than visual analysis and surveillance. The opposite applied to the specialist role. In addition, the nature of significant events differed in the two contexts. In general contexts, significant events tended to involve larger movements (e.g., a person crawling under a fence) that were easier to detect than the small, subtle and quick movements involved in specialist surveillance. The degree of specialisation varied according to factors such as industry, design of the CCTV system, and structure of the surveillance function.

Different observation skills and strategies are probably required for the two contexts. If this is the case, it is likely that work exposure and surveillance background could partly account for the high variance obtained and the lack of significant results. To get a rough estimate as to whether this could be the case, the backgrounds of participants who were previously excluded from the sample were examined. It emerged that two thirds (17) of the excluded participants performed a generalist surveillance function, suggesting tentatively that there may be validity in this distinction between surveillance backgrounds. It is reasonable to propose that participants with generalist versus specialist surveillance backgrounds perform differently on the video used in this research, which specifically reflected a specialist context. Divided on the basis of surveillance background, the sample included 13 generalists and 29 specialists. This is unbalanced, and resulted from the earlier exclusion of (mostly generalist) operators based on excessive FAs.

The biographical variables and SAMAE scores were analysed to ascertain whether any of these differed for generalists and specialists. T-tests on SAMAE scores were significant for the four- and fifteen-second accuracy scores on the Scanning test ($t=-3.19$, $p=0.003$ and $t=-2.16$, $p=0.05$ respectively), Scanning reaction time ($t=2.35$, $p=0.02$), reaction time and FAs for the Dynamic Attention test ($t=3.1$, $p=0.004$ and $t=2.32$, $p<0.05$ respectively), and the fifteen-second accuracy ranking score ($t=-3.52$, $p=0.04$). The only SAMAE measures that were not significant were FAs on the four-second measures for both Scanning and Dynamic Attention accuracy. Specialists mostly performed better than generalists on SAMAE.

Length of experience with CCTV surveillance differed significantly between surveillance backgrounds ($t=-2.73$, $p=0.01$), with specialists being more experienced but also having a greater standard deviation indicating a wider range of experience levels (Table 16).

Table 16. Summary statistics for CCTV experience for generalists and specialists (N=42)

Surv back	N	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Generalist	12	2.83	1.28	4.37	2.44	1.73	4.14
Specialist	29	5.85	4.09	7.61	4.62	3.67	6.25

Generalists and specialists had similar ages. However, they differed in terms of their education levels and training. While most (66.67%) generalists had less than a grade 12, most specialists (71.05%) had a grade 12 (Table 17).

Table 17. Analysis of education for generalists and specialists (N=42)

		Education		
		< Grade 12	**Grade 12**	**University**
Generalist	**Freq**	2	11	0
	Col %	66.67	28.95	0
Specialist	**Freq**	1	27	1
	Col %	33.33	71.05	100
Total	**Freq**	3	38	1
	Col %	100	100	100

The majority of generalists (83.33%) had only received on-the-job or technical training regarding operating the system (Table 18). This contrasted with 88.24% of the specialists who had been trained in surveillance skills.

Table 18. Analysis of training for generalists and specialists (N=42)

		Training			**Total**
Surv back		**Technical**	**Surveillance skills**	**Combination**	
Generalist	**Freq**	5	2	6	13
	Col %	83.33	11.76	31.58	30.95
Specialist	**Freq**	1	15	13	29
	Col %	16.67	88.24	68.42	69.04
Total	**Freq**	6	17	19	42
	Col %	100	100	100	100

A Chi-square test indicated a significant difference in the training received by generalists and specialists (Table 18), with $X^2(2)=10.63$, $p=0.005$. It is cautioned that some table cells had frequencies less than five. The Chi-square test is not necessarily valid when table cells have frequencies less than the minimum of five (Clarke & Cooke, 2004). Nevertheless, based on the above table it is evident that there were large differences in the types of training received by generalists and specialists.

There are consistent differences between participants from generalist and specialist surveillance backgrounds in terms of both SAMAE scores and biographical variables. However, these sub-samples were not compared on the dependent variables at this point in the analysis because such analyses were performed after additional data had been collected. Questions regarding the comparative detection performance, vigilance decrements and effects of the IGO intervention for generalists versus specialists are addressed in Stage 2 of the research.

7.1.3. Summary of results of Stage 1.

The level of detection performance (55% TB detection rate and high FAs) obtained by participants on the designated vigilance task was well below that of other CCTV studies where the significant event was known and visually salient, such as Wallace et al (n.d.) and van Voorthuijsen et al. (2005). This supports studies that suggest that vigilance performance requires improvement in the CCTV surveillance context (Cameron et al., 2008; Carli, 2008; Donald & Andrew, 2003; Donald et al., 2007; Gill & Spriggs, 2005; Keval & Sasse, 2006; Squires, 2010; Wells et al., 2006). However, the detection of TBs over time appears to have been statistically constant. This is contrary to the majority of vigilance studies that found a vigilance decrement (e.g., Craig, 1985; Hollenbeck et al., 1995; Mackworth, 1957; Molloy & Parasuraman, 1996; Parasuraman, 1979, 1986; Parasuraman et al., 1987; Pigeau et al., 1995; Schmidtke & Micko, 1964; Warm, 1984; Wiener, 1987). In addition, the intervention (IGOs – RIGOs and SIGOs) was not effective in terms of enhancing TB detection. These conclusions are highly tentative in view of the large variance in TBs and FAs (Table 11). Because of the large variance and unexpected findings, a decision was made to extend the sample and include participants with no prior experience with CCTV surveillance. The reasons for this are elaborated upon below.

High variance is typically caused by the nature of the sample and/or small sample size (Rosenthal & Rosnow, 2008). In terms of the sample, this study used operators who perform CCTV surveillance. This contrasts with the majority of vigilance research which uses student samples, such as Childs (1976), Cunningham, Scerbo, and Freeman (2000), Goldstein et al. (1969), Helton et al., (1999), Hollenbeck et al. (1995), Jerison (1959), Molloy and Parasuraman (1996), Rose et al. (2002), Sawin and Scerbo, (1993, 1994), and Wiener (1973). Operators and students are likely to have different levels of motivation and to approach the task differently, resulting in different vigilance dynamics. Similarly, it is possible that the intervention functions differently for participants with no experience in CCTV surveillance.

Extension of the sample to include students therefore meets the dual needs to examine the vigilance dynamics and the effects of the intervention on a different group of participants, and to increase the sample size.

The extension of the research to include students was also seen as a way of overcoming problems that were anticipated in terms of including more operators. The high rate of exclusion of operators in the current study meant that there was a strong chance that many operators would not meet the inclusion criteria, resulting in a high investment in time and cost compared with the value added. The exclusion of many operators is a key limitation of the current study. Although it may have been sensible to limit inclusion to specialists, remaining organisations with specialist surveillance were not accessible. Secondly, from an organisational perspective, participation disrupted shift arrangements and access to additional organisations was likely to be problematic. In addition, many organisations only employ a few operators, making access once again problematic.

The secondary analyses led to the identification of different types of CCTV surveillance, resulting in the distinction between generalists and specialists. It is emphasised that these terms are used to refer to the operators for convenience, but should more accurately be applied to their jobs as there were individual variations between operators within these groups. The allocation of participants to specialist or generalist sub-samples was not based on performance evaluations in their work settings, but merely on sort of surveillance conducted at their organisation. The terms 'specialist' and 'generalist' were chosen with care and contrast with those used in literature regarding expertise. Existing literature regarding expertise and skill acquisition refers to a progression through stages of expertise, such as novice, competence, proficiency, expertise and mastery (Dreyfus & Dreyfus, 1980). However, these frameworks apply to individuals and contrast with the current study where participants were categorised based on the nature of the surveillance function at their organisation and not on evaluations of on-the-job performance. Therefore the terms 'generalist' and 'specialist' are used rather than terminology derived from models of skill acquisition or expertise.

The differences between generalists and specialists in certain biographical variables and SAMAE scores suggest that surveillance background should possibly have been taken into account when allocating operators to experimental groups. However, this distinction was not initially apparent. The distinction between generalists and specialists could explain some of the variance in TB detection and may have impacted on the effect of the intervention. Therefore it appears that generalists and specialists should be treated as different strata in the sample based on their surveillance backgrounds, where surveillance background explains differences in vigilance performance and possibly, different effects of the intervention.

The division of the sample based on surveillance background provides an alternative way of looking at the data. Specifically, comparisons of novices (i.e., students with no experience in CCTV surveillance), generalists and specialists allows for the identification of potentially different trends for these sub-samples. The additional data and analyses are described in the next section, including the results of the qualitative analyses for both Stages 1 and 2.

7.2. Stage 2: Novices, Generalists and Specialists

This section reports on Stage 2 of the data collection and analyses incorporating data from Stages 1 and 2. Given the unexpected absence of a vigilance decrement in the first stage, similar detection rates for the experimental groups, the emerging importance of surveillance background and large variance in TB detection and FAs, a second stage of data collection and analysis was implemented. The use of an additional sample in the form of students or novices and the sub-division of the previous operator sample into generalists and specialists provide a different lens for examining the data. The aims of Stage 2 were to examine three emergent hypotheses which were variations on the hypotheses of the first stage, and allowed for comparisons between novices, generalists and specialists. After examining the findings for the hypotheses, responses to the interviews and other qualitative analyses for operators and novices are provided.

Previous research on expertise indicates that experts approach certain tasks (e.g., eye movements in driving; He et al., 2011) and use information differently from inexperienced job incumbents (Jarodzka, Scheiter, Gerjets, & von Gog, 2010). This suggests that participants with no background in surveillance are likely to approach the detection task differently from experienced participants and could respond to the intervention differently. In addition, IGOs could be used differently by novices compared with operators with surveillance experience. If this is the case, the intervention could have different effects for novices, generalists and specialists. This is supported by research on skill acquisition in a variety of contexts where experts perform various tasks better than novices. Examples of these tasks are driving (Deery, 1999), nursing (English, 2008), interpreting X-rays (Litchfield, Ball, Donovan, Manning, & Crawford, 2010), performing surgery (Law, Atkins, Kirkpatrich, & Lomax, 2004) and piloting aircraft (Kasarskis, Stehwien, Hickox, Aretz, & Wickens, 2001). Therefore it was decided to expand the sample to include novices. The inclusion of students as novices also allows for a better comparison with previous research that has used student samples.

The inclusion of novices leads to three emergent hypotheses. The first is that vigilance performance (TB detection, FAs, and accuracy) varies between participants with different surveillance backgrounds, as argued above. The second is that the vigilance decrement varies according to surveillance background. This is based on the dual findings of Stage 1 that a vigilance decrement did not occur for operators as a global sample, and that there was a large degree of variance between individual operators. The examination of the vigilance decrement in relation to surveillance background moves away from the assumption that all CCTV surveillance tasks are similar and provides an alternative way of looking at the data.

The third emergent hypothesis is that the intervention has different effects for observers with different surveillance backgrounds in the form of novices, generalists, and specialists. This is based on literature on expertise which reports differences in the higher order cognitive and perceptual processes that experts and novices adopt for complex, dynamic tasks (Deery, 1999). With tasks of this nature, experts perform better than novices at extracting relevant information from irrelevant information, making inferences from available data, are more flexible in their approach to the task, and use more short cuts (heuristics) based on knowledge (Jarodzka et al., 2010). These processes are required in CCTV surveillance, and could be influenced by IGOs via their potential influence on attention sets and mental schemas. Therefore IGOs could have different effects on novices, generalists and specialists.

7.2.1. Methodology.

The research design, materials, procedure and analysis methods were identical to those described in the first stage, with the exceptions of the need to gain access to the organisation, the separation of SAMAE and video sessions for most participants, and the additional analyses performed aimed at assessing the effects of different surveillance backgrounds.

7.2.1.1. Sample.

The additional sample obtained during Stage 2 consisted of 31 student novices who were new to the task of CCTV surveillance. Combining their data with those of Stage 1, this resulted in a total sample of 73. This creates a stratified sample (Rosnow & Rosenthal, 1996) consisting of 31 novices, 15 generalists, and 27 specialists. It allows for comparisons between participants with different surveillance backgrounds.

The novice sub-sample was obtained through convenience sampling (Rosnow & Rosenthal, 1996). The criteria for inclusion were that participants should be novices with no experience in CCTV surveillance, should not have an excessively high FA rate on SAMAE or the video. Thirty-three university students participated voluntarily in Stage 2. Two students were excluded based on high FA rates for SAMAE. No students reported high FAs for the video, yielding a usable sample of 31 (22 female, 9 male). Ages ranged from 20 to 28, with a mean of 21.55 years and a standard deviation of 1.48 years. All participants were studying towards Bachelors (19), Honours (10), or Masters (7) degrees. Nineteen participants were African, four were White, one was Coloured, and seven were Indian. These characteristics were spread fairly evenly between the experimental groups. As this sample consisted of novices with no experience in CCTV surveillance, the only other demographic variables recorded were age, gender, and year of study.

In addition to their lack of experience with CCTV surveillance, the novice sub-sample was younger, better educated, and included a greater proportion of females than the operator sample. In addition, it included Indian participants who were not represented in the operator sample.

7.2.1.2. Procedure.

As the procedure was the same as that for Stage 1 in most respects, only differences from Stage 1 are described. Permission was obtained from the relevant course coordinators and lecturers to request volunteers. Classes were told briefly about the study, participant information sheets were distributed and consent forms were signed by volunteers. The experiment was conducted at times convenient to participants. SAMAE and the video were completed on separate days to reduce the effects of fatigue. This is a departure from Stage 1 where most operators completed SAMAE and the video on the same day.

Allocation to treatment groups was based on SAMAE scores as in Stage 1 and summary statistics are presented in Table 19. Other information that was also used in the Stage 1 to allocation participants to experimental groups (e.g., experience, organisation) was not applicable to this sample and therefore only SAMAE scores were used. Where similar data for operators was presented in section 7.1, data is only presented for novices. To avoid confusion, the sub-samples included in analyses are noted in the titles of tables.

Table 19. Summary statistics of SAMAE scores per experimental group (Novices, N=31)

Group	Variable	Mean	SD	Min	Max
Benchmark (n=10)	Scanning accuracy 4 secs	44.29	12.10	23.81	64.29
	Scanning accuracy 15 secs	73.81	9.97	64.29	95.24
	Scanning reaction time	4.01	0.78	2.86	5.89
	Scanning FAs	2.80	1.55	1	6
	Dynamic attention accuracy	59.31	11.51	41.67	75
	Dynamic attention reaction time	1.86	0.15	1.64	2.05
	Dynamic attention FAs	8.10	7.64	0	21
	SAMAE 15 sec ranking	47.67	6.65	38.1	59.5
RIGO (n=11)	Scanning accuracy 4 secs	43.72	13.65	19.05	66.67
	Scanning accuracy 15 secs	71.43	12.46	45.24	88.1
	Scanning reaction time	4.05	0.91	2.87	6.14
	Scanning FAs	2.45	2.94	0	11
	Dynamic attention accuracy	55.81	10.86	40.28	77.78
	Dynamic attention reaction time	1.99	0.23	1.57	2.31
	Dynamic attention FAs	6.00	4.40	1	14
	SAMAE 15 sec ranking	47.52	4.40	40.1	56.2
SIGO (n=10)	Scanning accuracy 4 secs	42.86	13.37	19.05	57.14
	Scanning accuracy 15 secs	76.43	11.31	59.52	95.24
	Scanning reaction time	4.61	0.80	3.38	5.61
	Scanning FAs	4.10	5.38	0	18
	Dynamic attention accuracy	57.92	12.87	33.33	76.39
	Dynamic attention reaction time	2.09	0.39	1.62	2.93
	Dynamic attention FAs	12.00	16.18	0	54
	SAMAE 15 sec ranking	47.85	4.56	40.8	55.3
Total (N=31)	Scanning accuracy 4 secs	43.63	12.65	19.05	66.67
	Scanning accuracy 15 secs	73.81	11.15	45.24	95.24
	Scanning reaction time	4.22	0.85	2.86	6.14
	Scanning FAs	3.10	3.58	0	18
	Dynamic attention accuracy	57.62	11.44	33.33	77.78
	Dynamic attention reaction time	1.98	0.28	1.57	2.93
	Dynamic attention FAs	8.61	10.44	0	54
	SAMAE 15 sec ranking	47.67	5.10	38.1	59.5

The similarity between experimental groups on SAMAE scores was confirmed by ANOVAs which did not produce significant differences (Table 20). Therefore the groups were considered to be equivalent in terms of detection ability.

Table 20. ANOVAs of SAMAE scores for treatment groups (Novices, N=31)

Variable	Model df	Error df	F	p
Scanning accuracy 4 secs	2	28	0.03	0.97
Scanning accuracy 15 secs	2	28	0.51	0.61
Scanning reaction time	2	28	1.63	0.21
Scanning FAs	2	28	0.59	0.56
Dynamic attention accuracy 4 secs	2	28	0.24	0.79
Dynamic attention reaction time	2	28	1.77	0.19
Dynamic attention FAs	2	28	0.88	0.43
SAMAE ranking	2	28	0.01	0.99

* p<0.05, **p<0.01

7.2.1.3. Analyses.

Analyses were similar to those in Stage 1, with the addition of surveillance background as an independent variable and between-subjects measure. In addition, secondary analyses of biographical variables were not conducted as the inclusion of novices with no previous experience in CCTV surveillance and a different educational background would have skewed the results. Novices comprised a large proportion of the sample and had no previous experience in CCTV surveillance and had higher education levels than most of the operator samples.

7.2.2. Results.

In this section the results for Stage 2 are presented. First the results for the hypotheses are presented, followed by analyses of qualitative data. These include interview results and observation of non-verbal behaviour and periods of non-response that were interpreted as indicating task disengagement.

7.2.2.1. Examination of hypotheses.

Hypothesis 1: The level of detection performance differs between participants with different surveillance backgrounds.

Stage 1 established that the detection performance of operators on the designated task was lower than that in CCTV surveillance tasks where the significant event was known and visually salient, such as the studies by van Voorthuijsen et al. (2005) and Wallace et al (n.d.). Stage 2 examines whether this finding alters with a different sample, and the effect of surveillance background on detection performance (TBs, FAs, and accuracy). Means and standard deviations are reported in Table 21. Data for the experimental groups (benchmark, RIGO, and SIGO groups) was collapsed because no differences were found between groups, as reported later.

Table 21. Summary statistics for dependent variables by surveillance background (N=73)

	Surveillance background							
Variable	**Novice n=31**		**Generalist n=13**		**Specialist n=29**		**All participants N=73**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
TBs Phase A	3.10	1.54	3.46	1.27	3.52	1.45	3.33	1.45
TBs Phase B	2.10	1.45	2.15	0.90	3.17	1.17	2.53	1.34
TBs Phase C	2.48	1.55	2.38	0.96	4.14	1.41	3.12	1.62
Easy TBs	4.13	1.66	4.00	1.83	5.03	1.55	4.47	1.69
Difficult TBs	3.03	1.94	3.54	1.56	5.03	1.82	3.93	2.03
Total TBs	7.68	3.47	8.00	2.52	10.83	3.16	8.99	3.50
Total IGOs	22.38	1.99	19.33	3.74	21.68	2.11	21.55	2.62
FAs Phase A	2.61	3.02	10.00	8.36	5.24	3.62	4.97	5.25
FAs Phase B	2.13	2.03	5.77	4.32	3.83	4.30	3.45	3.72
FA Phase C	1.97	2.75	5.08	2.60	4.52	4.07	3.53	3.55
Total FAs	6.71	6.33	20.77	13.54	13.52	10.61	11.92	10.86

Vigilance performance for the global sample is discussed before that of participants with different surveillance backgrounds. The sample as a whole detected a mean of 8.99 (49.94%, i.e., 50%) TBs, with a standard deviation of 3.5. This indicates that on average, only half the TBs were detected and most participants detected between 5.49 (30.5%) and 12.49 (68.38%) TBs. Therefore a large proportion of TBs were missed. Inclusion of novices in the sample reduced the average percentage of TBs detected from 55% to 50%. The standard deviations for the dependent variables are high, especially for TBs in the last thirty minutes (1.62) and difficult TBs (2.03), indicating high variance amongst participants for these variables. Most participants in treatment groups detected the majority of IGOs. FAs were high with large standard deviations, indicating large variance between participants. Taken together, these results indicated that the detection accuracy of most participants for this vigilance task could be improved and that performance amongst participants was variable.

Results for surveillance backgrounds are now presented. Surveillance background had an effect on the mean TBs detected ($F(2, 65.4)=7.97$, $p=0.0008$). This applied to the least squares means of novices ($t=13.44$, $p<0.0001$), generalists ($t=8.5$, $p<0.0001$) and specialists (t=18.21, p<0.0001). Comparisons of the mean TBs found significant differences in TB detection between novices and specialists (t=-3.75, p=0.0004), and generalists and specialists (t=-2.72, p=0.0084) (Table 22). However, the mean TBs detected were similar for novices and generalists (t=-0.11, p=0.91).

Table 22. Differences in least square means for TBs by surveillance background (N=73)

Effect	Surveillance background	Surveillance background	Estimate	Std Error	DF	t Value	Pr > |t|
Surveillance background	Novice	Generalist	-0.04	0.36	65.4	-0.11	0.91
Surveillance background	Novice	Specialist	-1.03	0.28	65.4	-3.75**	0.0004
Surveillance background	Generalist	Specialist	-0.99	0.37	65.4	-2.72**	0.008

*p<0.05, **p<0.01

Inspection of the estimated means for TBs shows that specialists detected significantly more TBs (3.6) than both novices (2.57) and generalists (2.61). The trends regarding TB detection were analysed further in terms of easy and difficult TBs. Participants from all surveillance backgrounds detected similar number of easy TBs (F(2, 69)=2.87, p=0.06). However, the number of difficult TBs detected depended on surveillance background (F(2, 69)=9.18, p=0.0003). Tukey's HSD tests showed that specialists detected significantly more difficult TBs than both generalists and novices. The means for difficult TBs were similar for generalists and novices.

Surveillance background also accounted for differences in FAs (F(2, 65.9)=16.52, p<0.0001). This applied to novices (t=4.21, p<0.0001), generalists (t=9.27, p<0.0001), and specialists (t=8.34, p<0.0001). Comparisons of means between the surveillance backgrounds were all significant for FAs (novices and generalists, t=-5.65, p<0.0001; generalists and specialists, t=3.28, p=0.001; novices and specialists, t=-3.07, p=0.003) (Table 23).

Table 23. Differences in least squares means for FAs between surveillance backgrounds (N=73)

Effect	Surv back	Surv back	Estimate	Std Error	DF	t Value	Pr > |t|	Lower	Upper
Surv back	Nov	Gen	-5.65	1.001	65.9	-5.65**	<.0001	-7.65	-3.65
Surv back	Nov	Spec	-2.34	0.76	65.9	-3.07**	0.0031	-3.86	-0.82
Surv back	Gen	Spec	3.31	1.01	65.9	3.28**	0.0017	1.3	5.33

*p<0.05, **p<0.01

Inspection of the means for FAs indicates that novices reported the smallest number of FAs (estimated mean = 2.22), specialists an intermediate number (estimated mean = 4.57), and generalists the most (estimated mean = 7.88). The estimated means for total TBs detected and FAs reported for the different surveillance backgrounds are illustrated in figure 5.

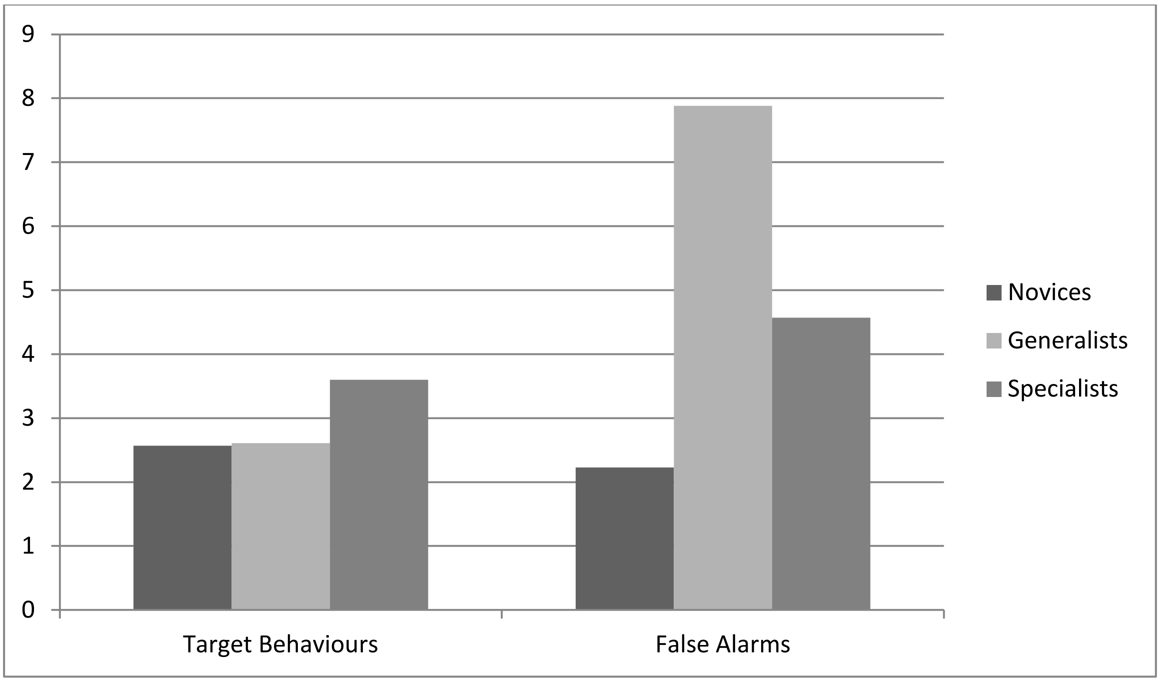

Figure 5. Total target behaviours and false alarms for novices, generalists and specialists

When TBs and FAs were combined into a normalised accuracy score, significant differences in means emerged for surveillance backgrounds (F(2, 69.2)=6.75, p=0.002). However, this applied only to generalists (t=-3.43, p=0.001). Comparisons of accuracy means between surveillance backgrounds are reported in (Table 24).

Table 24. Differences in least squares means for accuracy between surveillance backgrounds (N=73)

Effect	Surv back	Surv back	Estimate	Std Error	DF	t Value	Pr > |t|	Lower	Upper
Surv back	Nov	Gen	0.87	0.26	69.2	3.39**	0.001	0.36	1.37
Surv back	Nov	Spec	-0.02	0.19	69.2	-0.08	0.94	-0.40	0.37
Surv back	Gen	Spec	-0.88	0.256	69.2	-3.42**	0.001	-1.39	-0.37

*p<0.05, **p<0.01

Accuracy means for generalists were significantly different from novices (t=3.39, p=0.001) and specialists (t=-3.42, p=0.001). However, the means for accuracy for novices and specialists were similar (t=-0.08, p=0.94). Inspection of the estimated means shows that generalists obtained the lowest accuracy scores (-0.74), novices obtained an intermediate level (0.12) and specialists the highest scores (0.14). These results regarding accuracy for different surveillance backgrounds contrast with the findings for TBs where novices' mean TB detection rate was significantly lower than specialists'. However, novices' low rate of FAs increased accuracy to a level where it was similar to specialists. In addition, novices and generalists had similar mean TB detection rates, but the lower mean FAs reported by novices increased novices' accuracy to a point where their overall performance or accuracy was better than that of generalists.

Taken together, the results demonstrate that specialists performed best, novices performed at an intermediate level, and generalists performed worst on the task. The findings support the hypothesis that the level of detection performance varies between participants with different surveillance backgrounds.

Hypothesis 2: The vigilance decrement varies according to surveillance background.

To assess whether a vigilance decrement occurred for participants with different surveillance backgrounds, repeated measures ANOVAs are reported for thirty-minute periods and for each surveillance background. The detection difficulty level had a significant effect on TB detection for the global sample (F(1, 104)=11.76, p=0.0009). Because no significant differences were found between experimental groups (see hypothesis 3 later in this section) for TBs, FAs and accuracy, data for these groups were collapsed into one set for analyses regarding the vigilance decrement.

Surveillance background had a significant effect on TB detection (F(2, 65.4)=7.97, p=0.0008) and there was a significant interaction between surveillance background and phase (F(4, 126)=2.65, p=0.04). Separate analyses were computed to examine the effect of surveillance background on TB detection over time. Generalists and specialists (i.e., operators) were compared first, followed by separate analyses for novices, generalists and specialists.

To examine the vigilance dynamics in the operator sample from a different perspective from that in Stage 1, the mean TBs detected by generalists and specialists were compared. These comparisons found that the mean number of TBs detected was dependent on phase (F(2, 68.8)=7.13, p=0.002), surveillance background (F(1, 36.1)=1.49, p=0.005), and their interaction (F(2, 68.8)=4.3, p=0.02), all of which were statistically significant. This demonstrates that the mean number of TBs detected was affected by phase, whether operators were generalists or specialists, and the interaction between phase and status as generalist or specialist. Comparison of the mean TBs detected for different phases produced some significant results for the sample of operators (Table 25).

Table 25. Differences in least squares means for TBs by phase for generalists and specialists (N=42)

Effect	**Phase**	**Phase**	**Surv back**	**Surv back**	**Estimate**	**Std Error**	**DF**	**t Value**	**Pr > |t|**	**Lower**	**Upper**
Surv back			Gen	Spec	-0.99	0.3332	36.1	-2.98**	0.005	-1.67	-0.32
Phase	A	B			0.84	0.2431	65	3.44**	0.001	0.35	1.32
Phase	A	C			0.24	0.2909	101	0.81	0.42	-0.34	0.81
Phase	B	C			-0.6	0.2431	65	-2.47*	0.02	-1.09	-0.12

*p<0.05, **p<0.01

For the operator sample, ANOVAs on TB detection with surveillance background and phase as the independent variables found that the first and second phases (t=3.44, p=0.001) and second and third phases (t=-2.47, p=0.02) differed significantly on TBs. However, means for the first and last phases were similar (t=0.81, p=0.42). Examination of the estimated means for TBs shows that more TBs were detected in the first (3.47) than the second phase (2.63). The estimated mean for the third phase for operators is 3.23. These results show that there was a significant decrement between the first and second thirty-minute phases, and a significant increment from the second to third phases. This lends partial support to the existence of a vigilance decrement for generalists and specialists. These results differ from those in Stage 1 which found no significant changes in TB detection rates over time but which did not take surveillance background into account.

TB detection over time was next examined separately for novices, generalists and specialists to get a better idea of the vigilance dynamic for each sub-sample. For novices, the mean number of TBs detected differed significantly between phases (F(2, 54.7)=6.1, p=0.004). This applied to all three phases (Phase A, t=11.31, p<0.0001; Phase B, t=7.69, p<0.0001; and Phase C, t=9.06, p<0.0001). Comparisons of the means for TBs over phases for novices are presented in Table 26.

Table 26. Differences of least squares means for TBs for novices (N=31)

Effect	**Phase**	**Phase**	**Estimat e**	**SE**	**DF**	**t Value**	**Pr > |t|**	**Lower**	**Upper**
Phase	A	B	0.99	0.29	52.2	3.47**	0.0011	0.42	1.57
Phase	A	C	0.62	0.35	77.6	1.77	0.08	-0.08	1.31
Phase	B	C	-0.38	0.29	52.2	-1.31	0.2	-0.95	0.20

*p<0.05, **p<0.01

For novices, comparisons of the mean TBs only found a significant difference between the first and second phases (t=3.47, p=0.001). Inspection of the mean estimates shows that there is a decrement between phases A (3.11) and B (2.12). The mean for TBs increases for the last phase (2.4909) but this is not significantly different from the second phase (t=-1.31, p=0.2). The trend on TBs for novices is a decrement after the first half hour, followed by a flattening off of detection rates.

For generalists, phase was significant for TBs (F(2, 19.2)=8.67, p<0.002). This applied to all three phases whose estimated means were 3.42 for Phase A, 2.08 for Phase B and 2.33 for Phase C. Comparisons of the means between phases found significant differences between the first and second phases (t=4.16, p=0.0006) and first and last phases (t=2.78, p=-0.78, p=0.01). The second and last phases were similar for TBs (t=-0.78, p=0.45). The trend on TBs for generalists is that of a decrement which then flattens off.

Specialists showed a different trend regarding TBs over time. Phase made a difference to TB detection rate means (F(2, 49.5)=5.64, p=0.006). The means for TBs for the first phase were similar to those for the second (t=1.2, p=0.24) and third phases (t=-1.8, p=0.08), demonstrating that specialists were able to maintain their initial level of detection performance. Specialists' mean estimates were 3.52 for the first phase, 3.17 for the second phase, and 4.13 for the last phase. Unlike novices and generalists, the specialists' means on TBs were significantly different for the second and third phases (t=-3.35, p=0.002). Inspection of the means shows that for specialists, there was a significant increase in TBs detected in the last phase. The trend for specialists on TB detection was to maintain detection rates for the first hour, and then to increase detection in the last half hour.

The trends for TB detection over time varied with different surveillance backgrounds. These are illustrated in figure 6.

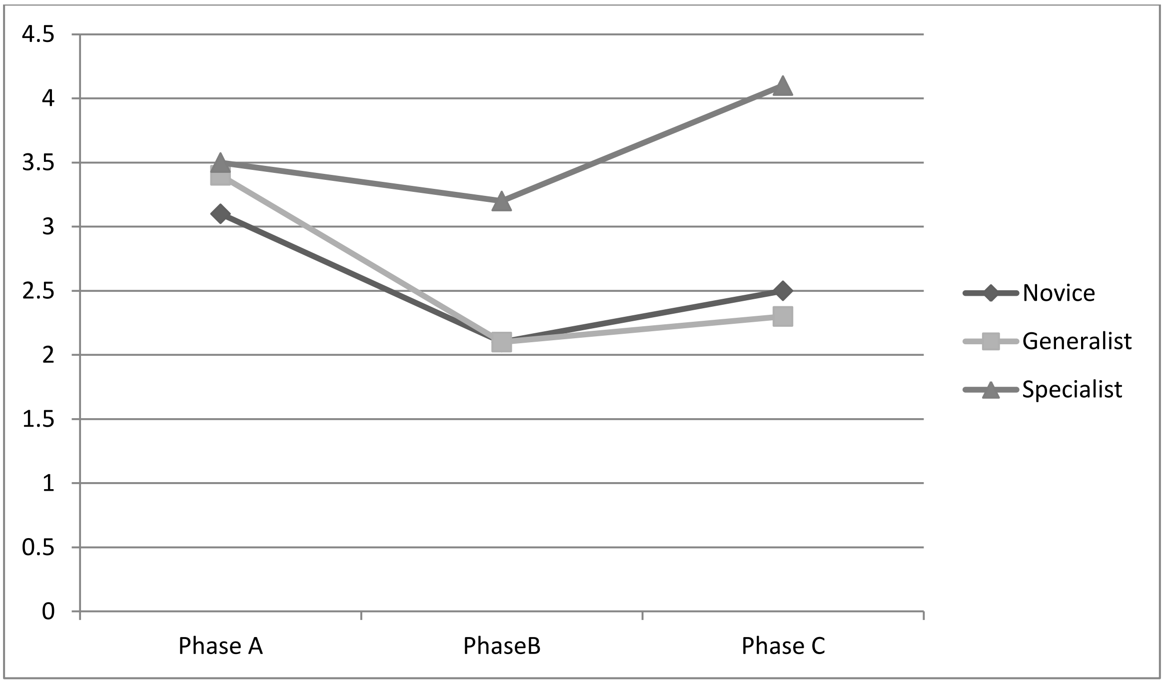

Figure 6. TB means over phases for novices, generalists, and specialists

The graph illustrates how specialists and generalists detected similar numbers of TBs in the first phase, and novices slightly fewer. The patterns for the three groups changed after the first thirty minutes. Specialists maintained their performance in the second phase, but novices and generalists decreased significantly. In the last phase, specialists significantly increased their performance and novices and generalists maintained their performance as their slight increment visible on the graph is not statistically significant. For TBs, specialists showed no decrement but increased their detection levels significantly towards the end of the task. By contrast, novices and generalists showed a significant decrement between the first two phases, after which their performance remained at the lower level.

The next dependent variable considered in relation to the vigilance decrement is FAs. FAs varied significantly depending on phase (F(2, 126)=17.45, p=0.0001) and surveillance background (F(2, 65.9)=16.52, p=0.0001) for the global sample. In addition, the interaction in the two-way ANOVA between surveillance background and phase was significant for FAs (F(4, 127)=6.10, p=0.0002). This was examined further through separate analyses for the different surveillance backgrounds.

The mean number of FAs reported by novices remained consistent across phases (F(2, 52.6)=0.64, p=0.53). The mean estimates on FAs for novices were 2.57 for Phase A, 2.12 for Phase B and 1.98 for Phase C. However, for generalists, the number of FAs reported changed between phases (F(2, 18.4)=8.87, p=0.002). The mean FAs reported differed significantly between the first and second phases (t=3.75, p=0.002), and first and third phases (t=3.83, p=0.0007). However, they were similar for the second and third phases (t=0.93, p=0.37). The mean estimate for Phase A was 11.58, Phase B was 6.64, and Phase C was 5.42. Inspection of these means indicated that for generalists there was a significant decrease in FAs after thirty minutes, after which the trend levelled off. For specialists, FAs differed between the phases (F(2, 52.8)=3.3, p=0.04). Specifically, FAs were significantly different between phases 1 and 2 (t=2.43, p=0.02) but were similar between phases 1 and 3 (t=0.96, p=0.3427) and phases 2 and 3 (t=-1.18, p=0.24). Inspection of the mean estimates of FAs for specialists indicates that there was a decrease between phases 1 (5.28) and 2 (3.86). The mean estimate for phase 3 was 4.55. The trend in FAs for specialists shows a significant decrease after thirty minutes, followed by a levelling off. These trends in FAs for participants with different surveillance backgrounds are illustrated in figure 7.

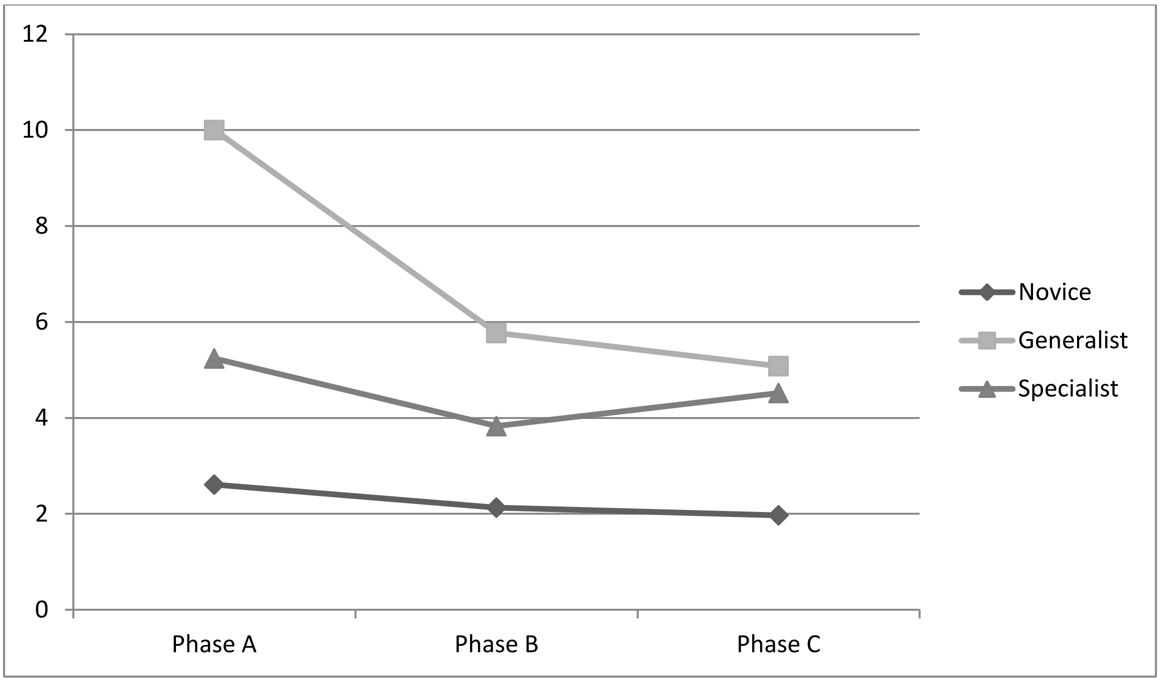

Figure 7. FA means over phases for novices, generalists, and specialists

The above figure illustrates the different trends in FAs for participants with different surveillance backgrounds. It is evident that novices consistently reported the fewest FAs and maintained a low FA rate over time. Generalists showed the highest rate of FAs initially, but this decreased significantly and then levelled off. Specialists reported an intermediate number of FAs at all phases compared with other participants. However, like generalists, FAs decreased significantly after thirty minutes. The slight increase shown by specialists in the last phase was not significant.

The overall accuracy measure consisted of a combination of normalised TB and FA scores. The TB difficulty index was related to accuracy for the global sample ($F(1, 111)=4.88$, $p=0.03$). No significant differences were found in accuracy levels between the phases for the global sample ($F(2, 128)=2.46$, $p=0.09$). Analyses for the sub-samples were conducted to determine whether differences between participants with different surveillance backgrounds were being masked in the global sample. However, accuracy did not change between phases for any of the sub-samples (novices: $F(2, 57.3)=2.94$, $p=0.06$; generalists: $F(2, 19.3)=1.54$, $p=0.24$; specialists: $F(2, 51.3)=2.29$, $p=0.11$). Accuracy was consistent for all sub-samples over the designated ninety-minute task.

The detection of IGOs was also similar over phases for the global sample (F(2, 86.6)=0.23, p=0.79). Once again, additional analyses were conducted for each sub-sample. These found that IGO detection did not change over time for novices (F(2, 39.6)=1.34, p=0.27), generalists (F(2, 13.3)=0.80, p=0.47), or specialists (F(2, 34.2)=0.66, p=0.52). This confirmed that IGO detection was consistent over time regardless of surveillance background.

The above analyses focus on differences between means and indicate trends between thirty-minute phases. However, the large standard deviations for TBs and FAs (Table 21) indicate a wide range of scores on the dependent variables. The standard deviations show variance for the time phases, but do not indicate variations in terms of when TBs are reported or missed within the thirty-minute phases. The patterns of hits and misses within phases were examined and examples of these are illustrated in Table 27 where 'X' indicates a correct detection for each participant. The full table is provided in Appendix 7.

Table 27. Examples of TBs detected over the ninety-minute task

	TB																	
	Phase A						**Phase B**						**Phase C**					
Obs	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
1	X	X				X	X	X	X							X		
2	X	X	X		X	X		X	X		X		X	X	X	X	X	
7		X	X					X									X	
9	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X
17		X			X	X			X		X		X		X			

Examination of the above table indicates that although some TBs were detected or missed by many participants, the patterns of detection varied considerably between participants. Two of these are illustrated with examples. Participant 1 detected the first two TBs, missed the next four, then detected four consecutively, missed six, detected one and then missed the last two. He presents a pattern with relatively long periods where no TBs are detected. This contrasts with participant 17 who missed the first but detected the second TB, then only detected TBs five, six, eight, nine, ten, fifteen and sixteen. This person appears to have frequent but relatively short periods where no TBs were detected. This analysis method is not intended to be rigourous, but merely illustrates individual differences in detection patterns that are not captured in the thirty-minute phase analyses. These patterns do not include an examination of whether FAs occurred during the periods when TBs were missed.

The absence of both correct detections and FAs could be interpreted as indicating fatigue and disengagement, as explained in the section on analyses. Estimates of the number of participants who showed at least one period of task disengagement in each time phase are provided in Table 28. These estimates were based on non-verbal cues and lack of responses.

Table 28. Disengagement by surveillance background and phase (N=73)

	Disengagement						**No disengagement**	
	Phase A		**Phase B**		**Phase C**		**All phases**	
	n	**Row %**	**n**	**Row %**	**n**	**Row %**	**N**	**Row %**
Novice (N=31)	12	38.71	21	67.74	20	64.52	7	22.58
Generalist (N=13)	2	15.38	10	76.92	7	22.58	2	15 .38
Specialist (N=29)	3	10.34	14	48.28	9	31.03	15	51.72
Total (N=73)	17	23.29	45	61.64	36	49.32	24	32.88

Almost a quarter (23.29%) of the sample lost concentration in the first thirty minutes, almost two thirds (61.64%) in the second thirty minutes, and about half (49.32%) in the last thirty minutes. In addition, novices (38.71%) tended to lose concentration sooner than generalists and specialists, although the majority of generalists disengaged in the second thirty minutes (76.92%). Specialists were not immune to disengagement, but smaller proportions of them lost concentration at all time phases than novices or generalists. Conversely, about a third of the whole sample (32.88%) showed no signs of disengaging during the video. Specialists had the largest proportion of participants who did not disengage (51.88%), followed by novices (22.58%) and generalists (15.38%). These trends are tentative as disengagement could be conflated with missed TBs and a lack of FAs. In addition, they were not accompanied by physiological or neurological measures of attention deployment. However, they provide an alternative perspective on the vigilance process.

In summary, there is partial support for a vigilance decrement, with TBs decreasing between the first and second thirty-minute phases and FAs steadily decreasing. However, the significant increase in TBs in the last thirty-minute phase is contrary to the vigilance decrement hypothesis. In addition, the high standard deviations for TBs in the last thirty minutes, difficult TBs and FAs indicate that dynamics regarding the vigilance decrement varied considerably between individuals. While differences were often related to surveillance background, there was also variance within surveillance backgrounds.

Hypothesis 3: The IGO intervention has different effects for observers with different surveillance backgrounds.

Means and standard deviations of the dependent variables for the experimental groups are reported in Table 29.

Table 29. Means and standard deviations for dependent variables for experimental groups (N=73)

	Benchmark (n=24)		**RIGO** (n=26)		**SIGO** (n=23)	
Variable	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
TBs Phase A	3.46	1.28	3.23	1.50	3.30	1.61
TBs Phase B	2.71	1.55	2.35	1.32	2.57	1.16
TBs Phase C	3.08	1.93	2.92	1.38	3.39	1.53
Total easy TBs	4.46	1.89	4.54	1.42	4.41	1.82
Total difficult TBs	4.13	1.87	3.38	2.10	4.36	2.06
Total TBs	9.25	3.65	8.50	3.15	9.26	3.79
Total IGOs	.	.	21.31	2.88	21.83	2.33
FAs Phase A	4.58	3.75	5.23	4.68	5.09	7.09
FAs Phase B	3.54	2.57	3.42	3.50	3.39	4.93
FA Phase C	3.88	3.08	3.50	3.75	3.22	3.87
Total FAs	11.96	7.84	12.15	10.49	11.61	14.01

Experimental group was not significant and had no effect on the detection of TBs ($F(2, 65.4)=0.09$, $p=0.91$), FAs ($F(2, 65.9)=2.18$, $p=0.12$), accuracy ($F(2, 69.2)=0.26$, $p=0.77$), or IGO detection ($F(1, 49.1)=0.31$, $p=0.58$) for the global sample. The interactions between experimental group and phase ($F(4, 126)=0.3$, $p=0.88$), and experimental group, surveillance background and phases ($F(2, 127)$-0.36, $p=0.94$) were also not significant for the global sample. This is similar to the results for Stage 1. The effects of the IGOs were also analysed for the sub-samples with different surveillance backgrounds as significant differences could have been masked in the global sample.

The mean number of TBs detected was consistent across experimental groups for novices ($F(2, 29.1)=0.97$, $p=0.3918$), generalists ($F(2, 10.1)=2.38$. $p=0.14$), and specialists ($F(2, 26.1)=0.14$, $p=0.88$). Therefore were no differences in TB detection between the benchmark, RIGO and SIGO groups for any participants, regardless of their surveillance background.

FAs were similar across experimental groups for novices (F(2, 26.7)=1.28, p=0.29) and specialists (F(2, 28.2)=0.77, p=0.47). However, FAs differed significantly across experimental groups for generalists (F(2, 9.42)=7.9, p=0.01). Comparisons of the mean FAs between the different experimental groups for generalists only are presented in Table 30.

Table 30. Differences of least squares means on FAs between experimental groups for generalists (N=13)

Effect	Group	Group	Estimate	SE	DF	t Value	Pr > |t|	Lower	Upper
Group	Benchmark	RIGO	1.64	2.02	9.42	0.81	0.44	-2.9	6.17
Group	Benchmark	SIGO	-7.03	2.39	9.42	-2.94	0.02*	-12.39	-1.66
Group	RIGO	SIGO	-8.67	2.21	9.42	-3.92	0.003**	-13.63	-3.7

*p<0.05, **p<0.01

For generalists, the mean number of FAs differed between the benchmark and SIGO groups (t=-2.94, p=0.02) and the RIGO and SIGO groups (t=-3.92, p=0.003). The mean estimate on FAs for the benchmark group was 6.08, for the RIGO group was 4.44, and the SIGO group was 13.11. Inspection of these means shows that the SIGO group reported significantly more FAs than the benchmark and RIGO groups. This suggests that SIGOs are likely to have significantly increased FAs for generalists only. The mean FAs for generalists were similar for the benchmark and RIGO groups (t=-3.92, p=0.003). These trends in mean estimated FAs for participants with different surveillance backgrounds within the different experimental groups are illustrated in figure 8.

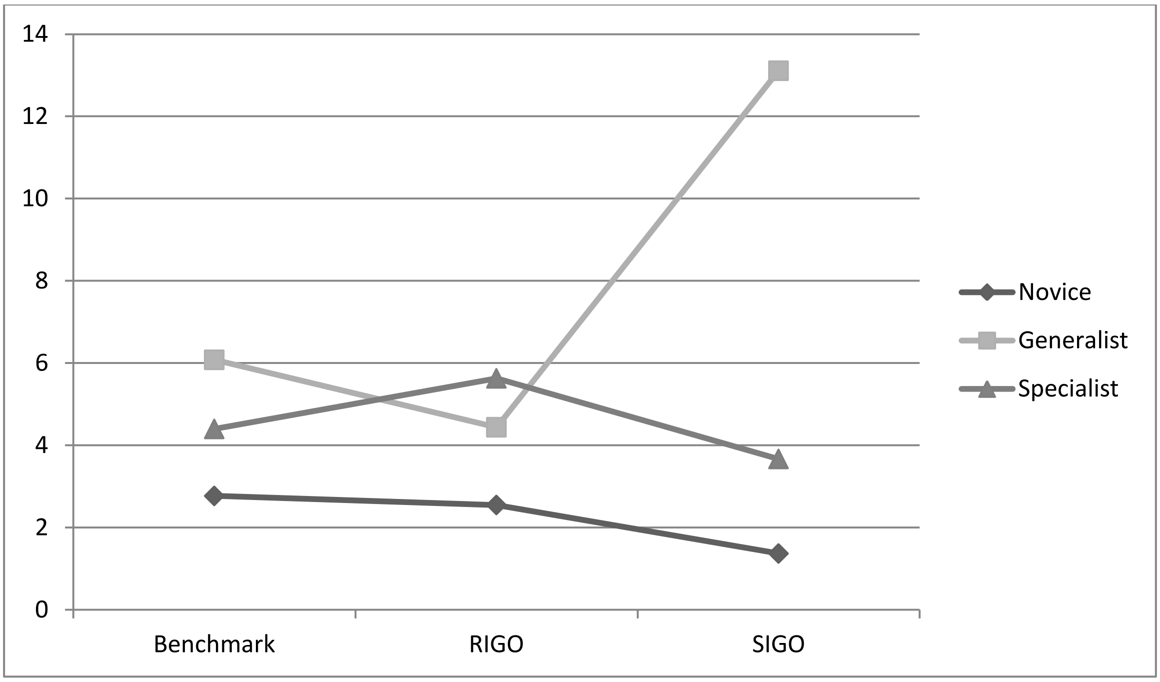

Figure 8. FA means by experimental groups for novices, generalists, and specialists

The above figure shows that FAs were fairly consistent across experimental groups for novices and specialists. However, the trend is very different for generalists with the mean FAs being significantly higher for the SIGO group compared with both the benchmark and RIGO groups.

Analyses on accuracy for the experimental groups for participants with different surveillance backgrounds were conducted. The experimental group had no effect on accuracy for novices ($F(2, 48.6)=0.96$, $p=0.39$), generalists ($F(2, 48.6)=0.96$, $p=0.04$), or specialists ($F(2, 43.2)=0.49$, $p=0.62$). Accuracy was consistent across the benchmark, RIGO and SIGO groups for novices, generalists and specialists.

The mean detection rates for RIGOs and SIGOs were compared for those groups for different surveillance backgrounds. The detection of RIGOs and SIGOs was similar for novices ($F(1, 23.1)=0.00$, $p=0.97$), generalists ($F(1, 7.13)=0.17$, $p=0.69$), or specialists ($F(1, 20.2)=0.06$, $p=0.81$). This suggests that RIGOs and SIGOs were comparable in terms of ease of detection for all surveillance backgrounds.

Pearson's correlations found no relationship between IGO and TB detection ($r=0.08$, $p>0.05$). Therefore IGO detection rates did not predict TB detection levels. Neither were any of the correlations between IGOs and TBs for separate experimental groups or surveillance backgrounds significant.

There was no support for the hypothesis that the intervention had different effects for observers with different surveillance backgrounds. The intervention did not affect vigilance performance for any of the experimental groups. In addition, no differences were found between the means of the RIGO and SIGO groups on any of the dependent variables, indicating that RIGOs and SIGOs were similar in their lack of effectiveness. However, qualitative analyses of observations and interviews identified some differences between RIGOs and SIGOs, as discussed in the next section.

7.2.2.2. Qualitative analyses: Reactions to the vigilance task and IGOs.

Participants' interview responses were analysed with the aim of exploring their attitudes towards the vigilance task and IGOs. This was intended to provide insight into the observation process from the participants' point of view.

Regarding the video, novices, generalists and specialists reported different attitudes to performing the vigilance task over an extended time and the video content. All novices reacted negatively to the observation task, regardless of which experimental group they were in (Table 31). Operators reacted more positively than novices, and showed greater variation between the experimental groups. In the benchmark group, more specialists (66.67%) were positive towards the vigilance task than generalists (14.29%) and novices (0%). In addition, more specialists were positive in the benchmark group than either of the treatment groups. This contrasted with the RIGO group where more generalists (60%) were positive about observing the video than specialists (37.5%) or novices (0%). A larger proportion of generalists reacted positively in the RIGO group than other groups. However, in the SIGO group, no novices or generalists made positive comments about the experience of watching the video, and only forty percent of specialists were positive about the video. Large proportions of participants expressed mixed reactions to the vigilance task in the benchmark (generalists 57.14%, specialists 33.33%) and SIGO groups (generalists 33.33%, specialists 20%), but not the RIGO group.

Table 31. Reactions to video (N=73)

	Benchmark			RIGO			SIGO		
Response	**Nov** (n=10)	**Gen** (n=4)	**Spec** (n=10)	**Nov** (n=11)	**Gen** (n=6)	**Spec** (n=9)	**Nov** (n=10)	**Gen** (n=3)	**Spec** (n=10)
Positive	0	14.29	66.67	0	60	37.5	0	0	40
Neutral	0	28.57	0	0	40	25	0	0	40
Mixed	0	57.14	33.33	0	0	0	0	33.33	20
Negative	100	0	0	100	0	37.5	100	66.67	0
Total	100	100	100	100	100	100	100	100	100

Positive responses to the video across experimental groups were that the task was interesting, challenging, and that participants learnt about significant events from it. Some participants (operators) also indicated a level of comfort and familiarity with the vigilance task because of its similarity to their surveillance function at work. Other participants who worked in a different type of environment from that depicted in the video reported that it was interesting because they learnt new things about a different environment. These operators approached the task as a challenge, using suspicious behaviours as cues for the specific TBs they were looking for. It is noted that responses referring to challenge and learning new detection skills were made mainly by specialists – the group that performed best at TB detection. Only two generalists but no novices made comments of this nature. Interview responses suggest that this group approached the vigilance task differently from the other sample groups. This is evident in terms of the search strategy they adopted and their motivation.

In terms of the search strategy, specialists demonstrated greater awareness of their search strategy than the other groups and may be assumed to adopt a different strategy from the other participants. This is reflected in several ways. Specialists applied their existing SA to the video and tried to develop it further. This is illustrated by the comment that the video was 'Interesting because I don't know that type of environment. I looked at body language and other cues, then you can pick up the TBs.' This quote also refers to the role of expectancies and attention sets in guiding attention. The importance of attention sets was elaborated upon by another specialist who said:

> 'If the person is working with a broom, you expect a broom flick and not a swivel. I saw drags [with foot] though sometimes. Your mind tells you what to

> look for, so you might miss things. You have a mind set about what is happening and you expect certain things. So you see these things. But it could be different.'

This illustrates the importance of attention sets being accurate as they could mislead attention if they are incorrect. One specialist commented that 'You have to know what to look for,' illustrating the central role of attention sets and SA in search strategy and suggesting that a goal-directed as opposed to random search strategy is seen as being most effective. The comments regarding attention sets illustrate that these specialists are aware of their attention set and their role in surveillance, suggesting meta-awareness of the observation process and self-monitoring. It is interesting to note that no novices or generalists made comments of this nature, suggesting that specialists were more aware of their strategies, employed different strategies, and/or were able to express their strategies more easily.

Specialists also emphasised the need for active analysis and interrogation of the display as part of their search strategy. This was apparent in comments that some specialists made while observing the video. These indicated that they used suspicious behaviours to form hypotheses in their minds which they then tested. During this process they adopted an analytical approach and showed a high degree of curiosity about what was occurring on screen. This is related to motivation and a strong desire to detect significant events. Quotes that illustrate their motivation are: 'You keep looking because you think something will happen' and 'You look more intensely because you think you are not getting enough.' Once again, novices and generalists did not make comments of this nature.

Participants from all surveillance backgrounds commented on the need to focus and concentrate and how this requires effort. However, responses indicate that novices and generalists found this more difficult and tiring than specialists. The time that it took for the TBs to occur made them difficult to detect. Specialists dealt with this by looking for suspicious behaviours and using these to guide their attention to the correct person and the TB: 'You have to be focused because things happen so fast. If you see suspicious behaviour before, it helps you to see the TB. You have to focus on the correct person.' They used SA to predict the occurrence of TBs, indicating a high level of SA. Once again, comments of this nature were not made by novices and generalists. Despite the emphasis on concentration, novices, generalists, and specialists commented that their minds wandered during the video, illustrating the difficulty of sustaining attention over time.

The interviews revealed a distinct pattern whereby specialists used a more active and goal-directed observation process. In addition, they were more aware of their attention sets and expectancies, as well as the potential for these to misguide them. They adopted a more complex view of the observation process than novices and generalists. All these aspects probably contributed to their more positive experience of the vigilance task and their ability to detect more TBs than novices and generalists. However, a substantial proportion of specialists in the RIGO group responded negatively to the vigilance task.

Negative comments regarding the detection task were that it was boring and tedious (for novices in particular), frustrating, too long, tiring and repetitive. Novices in particular used extremely negative adjectives such as ‘torture’ and ‘terrible.’ Frustration was related to the monotonous nature of the task, long periods between significant events and several sources of uncertainty. Uncertainty regarding the frequency of TB occurrence, which figure or area to observe, and the correctness of responses increased feelings of frustration. These are related to expectancies, spatial uncertainty and the need for feedback respectively.

Many participants reported that the detection task was difficult in terms of the nature of the detection task and the need to sustain concentration. The TBs were difficult to detect, especially when they did not occur in the foreground and when more than one person was in the video. However, more participants commented on the difficulty of maintaining concentration, as illustrated by the following quotes:

> ‘In the middle I got tired and lost concentration.’
> ‘Your mind starts to wander.’
> ‘You lose focus and become brain dead.’
> ‘I got distracted easily.’
> ‘My concentration went on and off.’
> ‘I felt like falling asleep.’

These quotes suggest fatigue, decreased arousal and depletion of attention resources. In addition, mind-wandering tended to remove focal attention from the detection task. This was particularly likely to occur when scenes were repeated and when the video showed little activity. Some participants reported difficulty in detecting TBs towards the end. This could indicate attention resource depletion or a sensitivity decrement over time. Although a statistically significant vigilance decrement was not found for all groups, a number of participants stated that 'It became too difficult to focus over time.'

Some participants who described the vigilance task in negative terms also suggested that they adopted a passive and disengaged search process. As one novice said, 'I was not engaging with the video, I was just staring.' Another described the task as 'Just sitting.' Some said they were 'Brain dead' during the video. This contrasts with the active search strategy adopted by many specialists. Some other participants, however, attempted to manage the deployment of attention resources: 'You try to concentrate but you feel tired, then you try to bring your mind back.'

The inability to sustain attention is related to many factors, including expectancies and the nature of the detection task. Novices were unique in their comments about the lack of action in the video. Their comments suggest that they associate videos with entertainment and that the CCTV video did not meet this need (e.g., 'There was no action,' 'Not challenging enough to maintain my attention. It was not interesting'). This is reasonable in view of the fact that they had no prior experience performing surveillance based on CCTV footage. Operators were more realistic in terms of their expectations.

Mixed reactions to the video included positive and negative comments. The positive comments referred to the video being interesting and learning new skills (e.g., 'The video gave light on seeing small things. You have to be more vigilant to see small things that a normal person wouldn't recognise. It was like a course – it helped me to see small details'). At the same time, however, these participants found it tiring and difficult to concentrate for the duration of the video.

In addition to obtaining general reactions to the vigilance task, the interview elicited reactions to the IGOs for the RIGO and SIGO groups (Table 32). Different reactions were noted to RIGOs and SIGOs between novices, generalists and specialists. While the majority of specialists (55.56%) were negative regarding RIGOs, many novices (45.45%) and generalists (28.57%) were positive. The pattern is different for SIGOs where most specialists (61.54%) and all generalists responded positively, but the majority of novices (60%) were undecided.

Table 32. Reactions to IGOs (N=73)

	RIGO			SIGO		
Response	**Novice** (n=11)	**Generalist** (n=6)	**Specialist** (n=9)	**Novice** (n=10)	**Generalist** (n=3)	**Specialist** (n=10)
Positive	45.45	28.57	33.33	30	100	61.54
Mixed	27.27	7.14	11.11	60	0	7.69
Negative	27.27	7.14	55.56	10	0	7.69
Total	100	100	100	100	100	100

Positive reactions to RIGOs focused on the perception that they increased alertness and therefore assisted concentration, as evident in the following quotes:

> ‘I didn’t look for them; just realised they were there. They could help operators because you don’t have to look for them and they just appear. They could make operators more alert.’
>
> ‘When focusing, you can pick up shapes very quickly. Shapes give you something to concentrate on. They give you more focus.’
>
> ‘After the shape popped up my attention would suddenly increase.’

RIGOs were described as a ‘wake-up call’ or reminder to concentrate. Their visual salience was seen as helpful as it made RIGOs easy to detect and increased interest and motivation. They were reported to reduce monotony and provide some relief when TBs occurred infrequently or were difficult to detect. However, the extent to which they increased alertness and the duration of this reported effect requires future research.

Negative comments about RIGOs focused mainly on their lack of relevance to the detection of TBs (i.e., arbitrary semantic distance) and potential to distract operators:

> ‘I can’t see how shapes could help operators … No point to them. I don’t see how they would help with concentration.’

> ‘Daft. They are obvious. They are separate from behaviours. They take your mind off the behaviours. We watch for people, not shapes.’

> ‘I think the shapes distracted more than anything. It forces you to focus on it, so you forget about the main behaviour.’

> “The shapes are easy to find. They made the things more interesting but didn’t help with the behaviour identification.”

Based on these comments, it is evident that participants saw RIGOs’ lack of relevance to significant events, or their arbitrary semantic distance, as being problematic. This induced feelings of annoyance in some participants, particularly some specialists. This was revealed in their tones when discussing RIGOs. This response to RIGOs could explain the high proportion of negative responses from specialists to the video (Table 31). Participants also emphasised their view that RIGOs do not assist with detecting significant events and related this to RIGOs’ arbitrary semantic distance.

Limitations to the potential benefits of RIGOs also emerged from the interviews. Although RIGOs were in some cases seen as improving alertness, some participants questioned the extent and duration of this effect. In addition, their potential to distract from significant events was seen as being problematic by some participants. Counter to this, however, a few participants mentioned that they missed some RIGOs because they were focusing on other areas at the time. This is consistent with the contingent involuntary orienting hypothesis (Folk et al., 1992). A few comments suggested that there was a trade-off between watching for RIGOs versus TBs. Some participants appear to have decided to prioritise the search for RIGOs and this could have had detrimental effects on the detection of significant events. A further disadvantage is that in operational contexts, RIGOs could become ‘part’ of the CCTV footage over time and lose their impact due to habituation. One generalist indicated that RIGOs increased mental workload by commenting that ‘There were too many things to watch for with shapes.’ Lastly, two generalists misunderstood the purpose of the RIGOs and interpreted them as accurate spatial cues. Their comments demonstrated that they believed that RIGOs indicated the area where they should look for TBs.

Operators responded more positively to the SIGOS than RIGOs, but novices expressed more mixed attitudes (Table 32). Most positive responses to SIGOs emphasised the SIGOs' role in creating an appropriate attention set through an understanding of what to look for in terms of the body language that accompanies the TBs. This is illustrated in the following quotes:

> 'The photos are helpful. I didn't realise they can help to focus more. The photos helped to look for something in particular. You need to know what to look for. With the photos, you try to see if things will be different next time. They help you to focus, not just look. At work, each day you look at the same things, but you have to see the differences. You get used to looking at the same things.'

> 'Pictures could help [operators] to see TBs on video because they show body language.'

> 'It [SIGO] shows you what to look for.'

SIGOs were perceived as enhancing SA and attention sets. Their ability to do this was related to the close relationship between SIGOs and the TBs they were intended to depict, or small semantic distances. These participants therefore endorsed the use of SIGOs with small semantic distances to significant events. Similarly to RIGOs, SIGOs were also perceived to enhance focus, concentration, and interest. However, the majority of positive comments regarding SIGOs mentioned their ability to assist with the detection of significant events through small semantic distances.

Only two participants made negative comments about SIGOs. One of these comments was that SIGOs would not work with live CCTV. Although the participant did not elaborate, s/he may have been alluding to various implementation issues. The second comment, from a novice, was that 'Pictures pulled my attention away from the video. They remind about the behaviours but don't help with identifying the TBs.' This comment is useful in distinguishing between the attention set and the ability to recognise significant events when they occur.

Mixed attitudes towards SIGOs referred mainly to distraction from TBs (3 participants). Despite this, these participants believed that SIGOs assisted the detection process by reminding them of the TBs and compensating for the potential distraction. In addition to distraction, several potential drawbacks of IGOs were mentioned. The first refers to SIGO design, with one participant stating that it was unclear which TB a few SIGOs represented. Another participant noted that the 'Pictures make you assume that the actions on the pictures were the same as those in the video. But those in the video were quite subtle and the behaviours in the pictures were easy to see.' In other words, recognising TBs in SIGOs was much easier than recognising actual TBs which occurred extremely quickly. As with RIGOs, some observers may interpret SIGOs as accurate cues for specific significant events and this could be detrimental to the detection task. Lastly, one operator noted that although SIGOs assisted with concentration, they did not reduce fatigue.

In summary, the IGOs were not effective in maintaining vigilance and motivation or enhancing the detection of significant events. Interviews suggest that SIGOs might have had a minor and short-lived influence on motivation, but this did not affect the detection of significant events. Lastly, there was no relationship between IGO and TB detection rates.

7.2.3. Summary of results.

Results for the global sample differed from those for specific surveillance backgrounds for the overall level of vigilance performance and changes in performance over time. Operators detected 55% of the TBs in the video and this decrease to 50% for the global sample when novices were included. Surveillance background had an influence on the number of TBs detected and FAs reported. Specialists detected significantly more TBs than novices and generalists. For FAs, novices reported the fewest and generalists, the most. Surveillance background was also significant for accuracy, with novices and specialists performing better than generalists. Although novices detected similar numbers of TBs to generalists, their significantly lower FA rate increased their accuracy to the point that it was similar to specialists. However, the high levels of boredom and frustration reported by novices suggest that this sub-sample would not be effective if they were employed in CCTV surveillance jobs over the medium to long term.

Results for vigilance decrements differed between Stages 1 and 2 of the research and were dependent on the sub-samples being analysed. Sub-samples either displayed a decrement or an increment, but not both. For novices and generalists, there was a vigilance decrement between the first and second phases after which performance reached a plateau. Specialists, on the other hand, maintained their TB detection rate for the first hour and then significantly increased it in the last thirty minutes when they anticipated that the end of the task was approaching. The vigilance dynamic over time for FAs also differed for participants with different surveillance backgrounds. Novices maintained consistently low numbers of FAs for the duration of the task, whereas generalists and specialists showed a decrease in FAs after thirty minutes, after which reports of FAs levelled off. Together, these results support the hypothesis that the vigilance decrement varies according to surveillance background. However, it is emphasised that a vigilance decrement did not occur for specialists.

The consistency in IGO detection rates over time is not surprising in view of their high salience. Previous research has found that the detection of less conspicuous, visually degraded stimuli decreases over time, but that this does not occur for highly salient stimuli (Neuchterlein, Parasuraman, & Jiang, 1983). As all IGOs were visually conspicuous and appeared abruptly, they were likely to capture attention.

IGOs did not have an effect on the detection of TBs, as evidenced by similarity in TB means between the benchmark, RIGO and SIGO groups. The absence of an effect on TB detection applied to all participants regardless of surveillance backgrounds. However, IGOs affected the number of FAs reported for generalists only, with generalists in the SIGO group reporting significantly more FAs than generalists in the benchmark and RIGO groups. This appears to be a negative effect that is specific to generalists. The results found no support for the hypothesis that the IGO intervention had different effects for observers with different surveillance backgrounds, or that it enhanced detection performance.

The next chapter discusses the results in relation to previous research and theory.

CHAPTER 8: DISCUSSION

This research contributes to a greater understanding of the nature of vigilance, the dynamics of the vigilance decrement, the applicability of these concepts in real world situations, and in particular the nature of the surveillance operator task and operations in a world where CCTV is becoming an increasing part of community and social protection. The findings provide a foundation for more effectively examining interventions aimed at improving detection levels and vigilance performance when dealing with complex, attention demanding, and dynamic tasks in a range of settings. The research terms contrast with those traditionally used in vigilance research, which is dominated by routine, monotonous and simple tasks with static visual stimuli. In this context it has important implications for how vigilance research can be approached and the ability to generalise findings to the real world. It is clear from this research that coping with real inputs, as represented by the behaviours in a CCTV viewing environment, makes the evaluation of vigilance dynamics a lot more complex than in studies with simple visual stimuli and tasks.

This chapter discusses the findings in relation to previous research, and identifies theoretical contributions of the current research. Three key themes are covered – vigilance performance, the vigilance decrement, and vigilance interventions, including the proposed IGO intervention. In addition, the theoretical implications and recommendations for organisations are mentioned where relevant.

8.1. Overall Level of Vigilance Performance

The research demonstrated a number of key findings regarding the overall level of vigilance performance. First, the rate of target behaviour (TB) detection was low, with a mean target detection rate of 50% for the global sample and slightly higher (55%) for operators. This contrasted with detection rates in the region of 90% in studies where the significant events were known and visually salient (van Voorthuijsen et al., 2005; Wallace et al., n.d.). Second, the task of detecting TBs that were visible, but at times performed quite quickly, was very difficult. This is indicated by the fact that no participants detected all TBs and only five operators (11.9%) detected 75% or more TBs. Third, there was a high rate of false alarms

(FAs), and differences in FAs were found between participants with different surveillance backgrounds. Fourth, the low detection rates and high FAs combined to create a low level of detection accuracy. Fifth, differences between participants with different surveillance backgrounds were identified in terms of detection rates, FAs, and their approach to the search process. Sixth, there was an appreciable variation in individual handling of vigilance demands and subsequent vigilance performance, as indicated in the statistical and qualitative analyses.

The low detection rates, high FAs and low overall accuracy indicate that operators had difficulty performing the task, despite the fact that all TBs were visible. Detection performance varies according to the task (Parasuraman, 1984) and therefore results of the current study should only be compared with those studies using a similar video detection task. Other studies have shown that even with simple detection tasks, detection performance does not achieve a perfect rate of one hundred percent (Wallace et al., n.d.). Studies on CCTV that achieved much higher detection rates than the current research used significant events that were known to participants and/or were visually salient (van Voorthuijsen et al., 2005; Wallace et al., n.d.). The reduction of uncertainty regarding the nature of significant events has a positive effect on vigilance performance (Parasuraman & Mouloua, 1996). Similarly, highly salient significant events tend to draw attention and are detected more efficiently during visual search (Pomplun, Reingold, & Shen, 2003). Therefore the high detection rates in the van Voorthuijsen et al. (2005) and Wallace et al. (n.d.) studies are not surprising. These contrast with the significant events (or TBs) used in the current research which were low in visual salience, occurred fleetingly, and where each TB took place through a number of different postures.

The only comparable study that uses a similar task and an operator sample is that of Andrew et al. (2003). These researchers also used a video consisting of CCTV footage from diamond processing plants but participants had to identify significant events involving theft. Their sample consisted of operators from a number of diamond mines in two countries. They found that detection rates for significant events varied between mines and countries, but obtained an average detection rate of 25 %. The detection rates of the current study are considerably higher than this, probably due to the fact that participants were required to detect specific TBs which were visible rather than significant events that required greater levels of interpretation of events and prediction of events. The differences in detection rates for CCTV surveillance tasks with different types of significant events suggests that CCTV surveillance is not a unitary task and that different skills and strategies are required for effective surveillance in different CCTV contexts. Further, it confirms the importance of operators' monitoring strategies (Wells et al., 2006).

Various studies show a decline in performance when viewing multiple CCTV monitors. A study on the relationship between the number of monitors viewed and observer performance used a surveillance task that required the detection of a designated person carrying an opened coloured umbrella in St Albans town centre (Wallace et al., n.d.). Observers viewing one, four, six, and nine monitors showed accuracy detection scores of 85%, 74%, 58% and 53% respectively in detecting the person with the umbrella. Wallace et al. (n.d.) also indicated that observers were significantly less likely to detect targets at greater depths (or in the background area of displays) within the image when there were more monitors. A separate study examining the effects of multiple monitors found that average detection scores were 91% for the single-monitor, 85% for the double-monitor, and 72% for the quad-monitor tests (van Voorthuijsen et al., 2005).

The low detection rate in this study suggests that expectations regarding CCTV detection rates require careful evaluation. A one hundred percent detection rate is unrealistic, even for top performers and those with extensive experience and training in CCTV surveillance. If this is the case, it is inevitable that significant events are missed in real operational contexts, although the proportion of missed significant events will vary according to the type of surveillance performed, the nature of significant events, and the skills and experience of operators. This finding indicates that it may be necessary to re-evaluate what is feasible in human detection and CCTV environments and the conditions that support and enhance detection. This is an important finding because it confirms that a greater emphasis should be placed on the operator's role in CCTV effectiveness and interventions aimed at improving detection rates. In addition, it highlights the importance of the overall level of vigilance performance which has been neglected by researchers. The establishment of an overall detection rate assists in evaluating whether performance is satisfactory, regardless of whether a vigilance decrement occurs or not.

These results appear to confirm the suggestions made in a number of studies that operator performance in CCTV is less than satisfactory (Donald & Andrew, 2003; Donald et al., 2007; Edkins & Pollock, 1997; Keval & Sasse, 2006; Wells et al., 2006), although the standards for 'satisfactory' detection have not been defined. Previous research has found mixed results for CCTV effectiveness (Greenberg & Roush, 2009; Keval & Sasse, 2006; Wells et al., 2006; Welsh & Farrington, 2002). However, none of these studies examined actual detection rates. Where the effectiveness of the overall CCTV system has been unsatisfactory, researchers have attributed this to a number of factors, many of which relate to the broader context in which the CCTV system operates, and not necessarily to operator performance (e.g., Cameron et al., 2008). By contrast, the current research focuses specifically on operator performance and demonstrates that operator detection performance is likely to be a significant contributor in situations where CCTV effectiveness is unsatisfactory.

Difficulties with CCTV surveillance and TB detection may be explained in terms of the complex nature of the CCTV surveillance task. The task included many of the characteristics associated with poor vigilance performance. These are, for example, the fact that it was a successive detection task (Parasuraman & Davies, 1977) with a crowded background (Bravo & Farid, 2006), and the significant events were complex (Koelega, 1996; See et al., 1995), low in salience (Mackie, 1987; Temple et al., 2000), uncertain (Folk et al., 1992; Parasuraman & Mouloua, 1996; See et al., 1995; Wolfe, 1998; Yantis & Jonides, 2000), and relatively infrequent (Koelega, 1996; van Wert et al., 2009) compared with many vigilance studies under experimental conditions.

Previous work exposure, in the form of surveillance background, had a major impact on participants' preparedness to detect significant events and their frame of reference for the task. None of the participants had previously been exposed to diamond processing plants, yet specialists developed an appropriate frame of reference for this context more readily than generalists or novices. They were better positioned to manage the recognition process and direct attention to relevant objects, events and behaviours, despite the fact that they had not previously been exposed to that environment. This potentially illustrates their higher level of surveillance skills and the ability to generalise their skills to different and new environments. This is consistent with skill acquisition processes which result from a combination of cognitive abilities, motivation, self-regulation, and task demands (Kanfer & Ackerman, 1996).

The high level of FAs reported by generalists and specialists reduces the accuracy of detection. More importantly, it has the potential to reduce the effectiveness of the system by demonstrating excessively high needs to respond to FAs and removes attention resources from real issues and events. The FA rate cannot be meaningfully compared with other research due to differences in the nature of the task and task length. Although Andrew et al. (2003) reported on FAs, the duration of their task differed from the current study and therefore direct comparisons cannot be made. FAs are attributed to a number of factors, including impulsivity (Helton, Kern, & Walker, 2009), guessing, inaccurate expectancies, a poor understanding of the TBs, poor visual analysis, and incorrect interpretation of visual stimuli (Endsley, 2000). The significantly lower FA rates for novices compared with generalists and specialists was surprising and could be explained by different expectancies between the student and operator samples, less desire than operators to demonstrate their detection skills (possibly a social desirability contaminant), and better visual analysis and decision making skills on the part of this particular sample of novices. Of these factors, the relationships between expectancies and vigilance performance, and decision-making processes in terms of the response criterion have been clearly demonstrated (MacMillan & Creelman, 1991). However, vigilance research has neglected visual analysis skills and aspects of decision making skills that are not related to expectancies.

Surveillance background emerged as an important variable influencing overall vigilance levels in terms of TB detection rates, FAs and accuracy. It was also related to participants' subjective experiences of watching the video and detecting IGOs. Participants from specialist CCTV surveillance operations detected significantly more TBs than generalists and novices (Tables 21 and 22). Previous vigilance research has not focused on group differences of this nature. However, the importance of surveillance background is consistent with a qualitative study where operators emphasised the importance of experience in CCTV detection (Tullio et al., 2010), although this study was not related in a quantitative manner to detection rates. The importance of surveillance background is supported by literature on expertise where experts performed better than novices on a variety of tasks, such as driving (Deery, 1999), flying aircraft (Kasarskis et al., 2001), performing surgery (Law et al., 2004), nursing (Benner, 1984), military control and command (Walker, Stanton, Salmon, Jenkins, Rafferty, & Ladva, 2010), and radiography (Litchfield et al., 2010) amongst others. This is related to higher levels of SA which assist decision making and job performance (Wright et al., 2004). SA is used to drive eye movements on a goal-directed basis (Di Stasi, Contreras, Cándido, Caňas, & Cantena, 2011; Hauland, 2008; Lorenz et al., 2006), increasing the likelihood that SA and eye movements are congruent with an effective viewing process. Surveillance experience and exposure play an important role in influencing how operators go about looking at the display, what they look for, when they do this, and where they look. This contributes to their frame of reference for the searching process.

The current study cannot be equated directly with research on expertise because participants were categorised based on their surveillance background and not on individual performance evaluations of expertise specifically. Despite this, there are similarities between the current study and research on expertise, as participants from specialist surveillance backgrounds typically had more experience and training, and presumably better skills than other participants. It is also possible that specialists and generalists went through different selection processes for their jobs, with specialist CCTV operations applying different, more stringent selection criteria than generalist surveillance operations. This could have contributed to different skill levels. This provides the first explanation for differences in detection rates between participants from different surveillance backgrounds.

A less obvious explanation of the higher detection rates for specialists is that they approached the viewing and detection process in a different manner from novices and generalists. This is based on the qualitative data indicating that specialists looked for different types of data from the other participants and interpreted them differently. It is consistent with eye movement research indicating that novices and experts look for and perceive different information (Jarodzka et al., 2010), resulting in different SA (Tien, Zheng, Swindells, & Atkins, 2010). In addition to searching for TBs, specialists looked for suspicious events that were associated with TBs and which could potentially 'cue' TBs. This provided information regarding locations or objects where they should direct their attention. This is consistent with research indicating that experts take different information into account from novices and apply it differently (Deery, 1999). Specialists demonstrated broader perception and comprehension of what was occurring in the video than other participants. This is supported by Tien et al. (2010) who found that novice surgeons tended to focus their eye movements and attention narrowly on the immediate task, thereby reducing awareness of highly significant events beyond the immediate focus on their attention. Differences in the breadth of focus of experts and novices is supported by numerous studies (Deery, 1999, Jarodzka et al., 2010; Kasarskis et al., 2001; Litchfield et al., 2010).

In addition to the type of information that specialists gathered, they also applied it differently from other participants. They used suspicious events to form hypotheses about what was occurring and then watched to see if they were correct. This process is not required in vigilance research using simple or meaningless stimuli and has several implications. Based on Endsley's (2000) model of SA, it illustrates the prediction of events based on a high level of SA and suggests that specialists had higher levels of SA than other groups of participants. The second implication is that specialists were motivated, interested and curious about what was happening in the video. This is captured in the comment that 'You look more intensely because you think you are not getting enough.' Thirdly, specialists visually analysed the display in a mindful manner. This suggests that they were engaged with the task and conducted an active search.

The distinction between active and passive searching suggests that successful CCTV surveillance is reliant on effective visual analysis. This contrasts with simple studies where the emphasis is on monitoring changes or differences in the physical properties of stimuli. The classic example of a study with simple stimuli is Mackworth's clock test where a double as opposed to a single click has to be detected (Mackworth, 1948). For the purpose of this research, visual analysis is defined as the active, goal-directed and effortful scrutiny of visual scenes in order to identify differences, discrepancies and patterns in objects, behaviours and events. Monitoring, on the other hand, is seen as a more passive process that involves watching for changes in objects, behaviours and events. It involves less in-depth scrutiny and analysis of the scene and is likely to draw on less SA. Visual analysis is seen as being highly intentional and goal-directed, while monitoring is likely to be more reactive and stimulus-driven. By adopting a perspective that distinguishes between monitoring and visual analysis, new areas of research are created.

The approach adopted by specialists implies that their searching was active, intentional, goal-directed, and effortful. In addition, specialists' active searching was supported by their motivation level, as indicated during interviews. Vigilance research acknowledges the importance of motivation (Davies & Parasuraman, 1982; Huey & Wickens, 1993; Mackie, 1987; Tomporowski & Tinsley, 1996) and effort (Smit et al., 2004), but says little about visual analysis, active searching, and how these processes interact with motivation. This active searching used by specialists contrasts with the passive approach adopted by novices and generalists. Many of the latter reported merely 'staring' at the display or 'just sitting,' indicating a passive, 'wait and see' approach. Several reasons are proposed for this. The task may have required high levels of attention resources and effort, especially for novices. This is consistent with the multiple resource theories of attention (Polson & Friedman, 1988; Boles, 2002; Boles & Law, 1998; Wickens, 2002) and the different amounts of cognitive resources and effort associated with different phases of skill acquisition (Kanfer & Ackerman, 1996). Expectancies regarding the nature of the task are likely to have differed between novices and operators. Novices had no previous exposure to CCTV surveillance and probably associated the concept of watching a 'video' with passive entertainment rather than with an attentionally draining task. These factors would influence motivation and expectancies regarding the nature of the task, which in turn could have affected performance. However, this does not explain why the generalists' results were similar to the novices. In view of the different educational backgrounds of the generalists and novices, it is possible that the generalists had poorer visual analysis skills than the novices and that these resources were drained quickly, leading to a more passive searching process.

Novices experienced greater task disengagement than operators. Many novices stated that they had difficulty controlling mind-wandering during the task. This is supported by the earlier onset of motor indicators of fatigue and disengagement in novices and generalists compared with specialists (Table 28). These processes are consistent with the relationship between disengagement, task unrelated thoughts and task-related interference (Smallwood et al., 2004). A mindful approach to surveillance may be adopted initially, creating attention resource depletion and task disengagement, whereupon further observation becomes passive until re-engagement occurs. However, this did not occur for specialists who were more able to maintain their attention and even increase their detection rates towards the end of the task.

The link between active search and better detection performance strongly suggests that active, intentional and mindful search is required for effective CCTV detection. Active search is essential to CCTV surveillance whereby operators intentionally form and test hypotheses about the scenes, apply high levels of SA in order to predict significant events, and intentionally regulate their attention resources and task engagement. This tentatively suggests that operators who are best at detection monitor themselves through meta-cognitive strategies where they are aware of the approaches they use, monitor and evaluate these, and are able to alter them as required. This is consistent with Endsley's (2000) emphasis on the use of meta-cognitive strategies in the application of high levels of SA. This may have implications for other work tasks requiring substantial visual analysis, such as baggage screeners and air traffic controllers.

8.2. Vigilance Decrements

Having discussed the overall level of detection performance, the discussion moves to results concerning the vigilance decrement, or changes in vigilance performance over time. Regarding the vigilance decrement, the current research found first, that a clear and statistically significant vigilance decrement did not occur in a linear manner over the designated ninety minute time period for all participants. A statistically significant decrease in TB detection rates occurred between the first and second thirty-minute periods for novices and generalists but not for specialists. Second, specialists maintained their detection rate for the first hour and then increased it significantly in the last thirty minutes. This is inconsistent with the generally expected vigilance decrement in the literature (Pigeau et al., 1995). Third, individual decreases in detection performance occurred at different times within the ninety-minute period for different participants. This shows a high degree of individualised responses and variations in strategies for coping with vigilance demands. It is consistent with individual variability in attention (Conway & Kane, 2001; Tomporowski & Tinsley, 1996). Fourth, behavioural indices of task disengagement as measure by the research administrators were associated with detection failures. Fifth, group differences based on surveillance background or work exposure played a major role in changes in vigilance performance over time. Some of these findings contrast with the majority of previous vigilance research which implies that a vigilance decrement will occur, that this will happen within 20 to 35 minutes or sooner (Neuchterlein et al., 1983; Sawin & Scerbo, 1995; Robertson et al., 1997), and that it occurs in a linear manner (Pigeau et al., 1995).

This research highlights the nature of differences in complex tasks based on work in the real world as opposed to the simplistic, routinised and highly repetitive tasks often used in laboratory experiments. The absence of a single, linear vigilance decrement for all participants and the specialists' ability to maintain and then increase performance, contrasts with the majority of vigilance research conducted in laboratories with simple stimuli (e.g., Craig, 1985; Mackworth, 1957; Parasuraman, 1979, 1986; Parasuraman et al., 1987; Schmidtke & Micko, 1964; Warm, 1984; Wiener, 1987). However, research on complex and real life vigilance tasks has obtained mixed findings regarding the vigilance decrement, with certain studies reporting a decrement in detection rates (Hollenbeck et al., 1995; Molloy & Parasuraman, 1996; Pigeau et al., 1995; Wiggins, 2011) and others not finding a decrement (Adams et al., 1961; Moray & Haudegond, 1998). A number of explanations for these differences in findings have been proposed, such as the presence of moderating influences (Davies & Parasuraman, 1982) and differences regarding the types of tasks used (Helton et al., 2010; Hollenbeck et al., 1995), individual differences and the types of samples used. Realistic tasks involve more interesting, complex and dynamic stimuli and environments than most laboratory experiments (Donald, F., 2010). Although some participants experienced the task as boring (particularly the novices who would represent samples most often found in laboratory studies), the nature of the scenes depicted was probably more interesting than the simple stimuli used in most laboratory studies and may therefore have been more capable of holding participants' attention. The content of the scenes is imbued with meaning (Donald, F., 2008), particularly for operators who are accustomed to observing scenes.

The findings of the current research strongly suggest that the nature of the sample is a significant and frequently overlooked factor in relation to the vigilance decrement. A vigilance decrement was found for the novices (i.e., students). Closer examination of the studies with complex tasks referred to above, indicates that a number of them used student samples despite using tasks or simulations that were similar to those encountered in the real world (e.g., Hollenbeck et al., 1995; Molloy & Parasuraman, 1996; Pigeau et al., 1995). However, these studies differ from the current research in that they did not compare the vigilance dynamics of students with those of experienced job incumbents. The current research found a distinct difference in the vigilance dynamics between participants (novices and generalists) with no or less experience, training, and exposure to the type of surveillance task performed in the study, and specialists with more experience, training, and exposure to this type of task. This demonstrates that previous work exposure is an important variable to consider when selecting a sample for research, and that the results of studies that use student samples on complex tasks may not apply to experienced job incumbents. For operators, the video content and observation process resonated with their daily work activities. However, the task held less meaning for novices, none of whom had previously attempted CCTV surveillance. This suggests that research into vigilance tasks with a bearing on real world jobs should be conducted on samples with relevant work exposure. The involvement of samples lacking appropriate work exposure is likely to reduce ecological validity. Conversely, the results of simple, meaningless tasks with little relation to real world vigilance tasks may not generalise to all personnel in operational contexts. The current research calls into question the ecological validity of research on students, especially on tasks that are intended to simulate jobs in the real world.

The similar vigilance dynamics for novices and generalists suggest tentatively that people with little or no exposure to particular vigilance tasks are likely to be more vulnerable to the vigilance decrement compared with people with relevant and extensive work exposure. However, this is contrary to a number of vigilance studies that found a steady vigilance decrement for both novice personnel and experienced personnel (Baker, 1962; Mackworth, 1970; Pigeau et al., 1995; Schmidke, 1976). In these studies, both novices and experts were employed to perform the task used in the relevant study. This differs from the current research where novices were not operators and were drawn from a different population from operators with different expectancies, motivation and skills.

The third, fourth and fifth key findings regarding the vigilance decrement concern the form of changes in detection performance over time. For novices and generalists (but not specialists), the significant decrease in TB detections was followed by a levelling off of TBs detections. This is consistent with the shape of the vigilance decrement curve that has been consistently reported in the past with simple tasks (Pigeau et al., 1995). In addition to this, different patterns of hits and misses were found between individuals within the phases. Previous research with complex tasks has seldom examined what happens in terms of patterns of hits and misses on a moment-to-moment basis within the vigilance decrement curve, or in situations where no vigilance decrement occurred (with the exception of Howell et al. (1966) who found alternating lapses in attention rather than a vigilance decrement). Moment-by-moment fluctuations in performance (Gilden & Hancock, 2007) have been examined in certain simple tasks, such as the SART (Robertson et al., 1997), but these tasks are very different from the complex video task used in this research. The likely reason for this neglect is an assumption that the vigilance decrement curve is the most important vigilance performance issue, as suggested by the predominance of theories and research that focus on this form of decrement. The current study found, however, that the vigilance decrement curve does not apply to all groups of participants, that certain participants achieved the opposite trend of detection maintenance followed by a performance increment, and that variations in patterns of hits and misses occurred between individuals. This is consistent with researchers who have found individual variability in attention (Derryberry & Reed, 2001; Rose et al., 2002; Tomporowski & Tinsley, 1996).

The individualised patterns of hits and misses and the absence of a vigilance decrement for specialists suggest that the terms ‘performance fluctuations’ and ‘vigilance dynamics’ are more appropriate than ‘vigilance decrement’ for complex vigilance tasks. These terms allow for moment-by-moment changes in attention deployment and task engagement, the maintenance of detection levels over time, and performance increments and decrements. They are more inclusive of various types of performance changes without assuming that a vigilance decrement is the only or most important issue in vigilance.

Task engagement emerged as a key process associated with performance fluctuations. This is based on the finding that misses were often accompanied by task disengagement. It is supported by studies reporting a relationship between task disengagement, attention failures, and decreased performance (Cheyne et al., 2009; Dockree et al., 2005; Gilden & Hancock, 2007; Grier et al. 2003; He et al., 2011; Manly et al., 1999; McVay, Kane, & Kwapil, 2009). However, it is noted that task disengagement was not the only reason for poor or decreased performance and other explanations are also discussed in this chapter. Periods of task disengagement varied considerably between individuals in terms of onset time, duration and frequency (Table 28). Most (67.17%) participants disengaged from the video at various times, suggesting that task disengagement was highly pervasive. The starting time of the first disengagement period varied considerably between individuals, with some disengaging within the first thirty minutes but most during the second thirty minutes (Table 27). In addition, most participants demonstrated alternating periods of engagement and disengagement, with the length and frequency of disengagement periods varying between individuals. This alternating pattern is consistent with that reported by Cheyne et al. (2009) in a study on task disengagement. The alternation of periods of task engagement with attention 'time outs' assists the preservation of attention resources over time (MacLean et al., 2009, p. 23). The need to take 'time outs' indicates that most participants in the current research experienced difficulty in sustaining attention which impacted on their ability to maintain concentration and task goals.

Patterns regarding disengagement also show different trends for sub-samples with different work exposure (Table 28). Novices tended to visibly disengage from the task sooner than generalists and specialists, and specialists had the smallest proportion of people who visibly disengaged in the first thirty minutes. These results indicate that specialists were best and novices worst at regulating and maintaining attention. A corresponding result is that novices reported boredom most frequently. It is possible that one of the skills associated with different types of surveillance background is the ability to self-regulate attention, and that this reduces the vigilance decrement. However, it is not known whether specialists were initially selected for their jobs because of their superior abilities to regulate and sustain attention and detect significant events, or whether they had developed this over time with work experience and training. The development of sustained attention has been examined from a developmental perspective (Levy, 1980) and in terms of disorders such as attention deficit disorder (Barkley, 1997; Hollingsworth, McAuliffe, & Knowlton, 2011), but has been neglected by researchers in relation to real world, complex vigilance tasks performed by adults.

The content of the video was often, but not always, related to task disengagement. Disengagement and mind-wandering frequently occurred during scenes with little activity or high repetition, especially if these scenes were perceived as lasting for extended periods. Participants interpreted these scenes as containing little information of interest or relevance to the detection task. This is consistent with the extended, monotonous, repetitive and familiar conditions under which task disengagement typically occurs (Cheyne et al., 2009).

Previous research links boredom and disengagement (Donald, 2001) but has not examined the link between boredom and work exposure. Boredom was highly pervasive in the novice sub-sample. Some researchers associate boredom with a mindless approach to vigilance (Robertson et al., 1997) but others link it to attention resource depletion (Grier et al., 2003; Helton et al., 2010). In the current research it is unclear whether a mindless approach resulted in disengagement and boredom, or whether the depletion of attention resources caused difficulties in sustaining attention and boredom. However, previous research has linked novice status to higher resource depletion (Kanfer & Ackerman, 1996), suggesting that in the current research the attention resources of novices were likely to be depleted more quickly than those of operators. This requires further research.

Three sources of evidence emerged demonstrating that participants attempted to regulate their attention during the video. First, participants reported efforts to maintain concentration. Second, the alternating periods of engagement and disengagement suggest a degree of attention management. Last, the significant increase in performance in the last thirty minutes of the video for specialists indicates that these participants were able to re-engage during that period. Taken together, these suggest that a mindful approach and subsequent resource depletion as the underlying mechanism for disengagement rather than mindlessness. This tends to support the mindful (Grier at al., 2003) rather than the mindlessness theory of vigilance (Robertson et al., 1997).

The increase in performance towards the end of the task for specialists deserves further mention. Because participants were aware of the duration of the task, the increase may have been based on knowledge that the task was coming to an end, indicating voluntary, endogenous attention processes. This is consistent with the end spurt effect described by Bergum and Lehr (1963). In the current research, the increase has several implications. It shows that certain participants intentionally regulate their task engagement and attention resources in certain situations and are successful in this to some degree. This is consistent with Maclean et al. (2009) who found that observers regulate attention resources to a certain extent. The specialists' increase in performance demonstrates that in some instances, operators are able to deliberately overcome issues related to arousal, fatigue and alertness when they believe that the end of the vigilance task is near. This contributes to the maintenance of performance levels and prevention or reduction of a vigilance decrement. Self-regulation implies that the effects of fatigue and other processes described in vigilance theories are to some extent reversible. Future research is needed regarding the mechanisms and contexts for re-engagement.

The focus on the self-regulation of task engagement indicates a different emphasis and vigilance dynamic from the majority of research on the vigilance decrement. As indicated, much vigilance theory is based on a steady vigilance decrement and the need to sustain attention in a linear manner on a continuous basis. Although constant, continual attention is the ideal, it appears to be unrealistic for complex sustained tasks. The recognition of fluctuating levels of task engagement changes the focus of research and theory to attentional and other cognitive mechanisms (e.g., motivation) involved in task disengagement and re-engagement. Further, it emphasises the need to identify ways of renewing motivation and interest in the task and enhancing the self-regulation of task engagement. Although both approaches deal with attention, they approach the issue from different angles.

Task differences create different frames of reference for observers and understandings of the task which are likely to impact on task significance, motivation, and performance. In addition, information processing and other cognitive demands must inevitably differ between complex, dynamic and simple, static tasks (Parasuraman & Davies, 1977). These in turn are likely to influence how participants approach the task, attention sets (Leber & Egeth, 2006; Most et al., 2005), eye movements (Castelhano & Henderson, 2008), and the use of exogenous and endogenous attention guidance (Serences et al., 2005) in searching for significant events.

The notion that the vigilance decrement may not be as prevalent, linear and universally applicable as implied in much vigilance research, suggests that a shift in the focus of research towards the overall level of vigilance and related attention processes is needed. However, this requires a better understanding of the processes involved in the overall level of vigilance and their interrelationships, including the role of task engagement and attention fluctuations.

Differences between generalist and specialist CCTV contexts suggest that CCTV surveillance is not a unitary function and that the nature of the work and related demands on operators vary across contexts. This is an important distinction as it represents the first step in developing a typology of CCTV surveillance jobs and the cognitive processes involved. Previous research has assumed that all CCTV surveillance contexts are similar in terms of demands on operators. Differences between contexts have implications for theory and practice. The identification of CCTV contexts allows for future research to be specific in terms of the contexts to which the research is intended to apply. It also has implications for the selection, training and development of operators. Existing typologies of vigilance tasks (e.g., Parasuraman & Davies, 1977; Donald, F., 2008) could be used as a starting point for a typology for CCTV surveillance tasks.

The original contributions made by the current research in terms of levels of vigilance performance and fluctuations in performance levels over time are summarised as follows.

- The overall vigilance performance level is extremely important in CCTV surveillance and probably in other complex, real life vigilance tasks as well. For some operators, it may have a greater effect on detection rates than the vigilance decrement.
- Work background is a variable that has not previously been examined. Past vigilance research has given inadequate consideration to the type of work exposure of research samples and differences in findings for different samples.

- Fluctuations in attention and task engagement on a moment-by-moment basis contribute to attention and detection failures.
- The nature of the task content is associated with engagement. This has not previously been examined in the CCTV context.
- The approach to the searching process differs for samples with different work exposure.
- The identification of passive versus active searching strategies in CCTV surveillance is seen as being crucial to successful detection. Active and passive search are related to factors such as motivation, attention self-regulation and task engagement.

8.3. The IGO Intervention

Having discussed the overall detection performance and changes in detection over time, established the need for an intervention and examined certain of the attention processes involved in searching video displays, the discussion moves on to the IGO intervention. The main purpose of the research was to develop an IGO intervention to enhance the detection levels of operators involved in proactive or real time CCTV surveillance. IGOs were aimed at maintaining vigilance and motivation, assessing whether operators are looking at the display in order to maintain significant event detection rates, and enhancing the detection of significant events. The key findings regarding the IGO intervention are as follows. First, there were no significant differences between the benchmark, RIGO or SIGO groups in terms of TBs, FAs, or accuracy. Therefore the IGOs did not influence these outcomes and were not effective in enhancing the detection of significant events. However, there were small differences in performance between the experimental groups. The group that received SIGOs showed the largest improvement towards the end of the experiment. Although this was not statistically significant, it suggests that the IGO intervention has the potential to assist the detection process if developed further, or if the detection task were longer. Second, RIGOs and SIGOs did not differ in terms of their lack of effect on TB detection, but there are indications that they had different but small effects on the attention processes of participants. This is based on different affective reactions to RIGOs and SIGOs, and comments regarding distraction by RIGOs in particular and the usefulness of SIGOs in relation to attention sets. Third, the high visual salience of all IGOs was perceived by participants as an advantage in that IGOs were easy to detect. However, this did not translate into improved TB detection rates. Fourth, the images depicted in SIGOs which all had small semantic distances to TBs were perceived more positively than the random images depicted in RIGOs which were unrelated to the detection task.

The fifth key finding regarding the IGO intervention was that the potential impact of the IGOs on TB detection was the same across participants from different surveillance backgrounds. Therefore the effectiveness of the IGOs did not vary across surveillance backgrounds. Sixth, IGOs did not influence the detection of TBs in sub-samples where a vigilance decrement did occur. Seventh, only generalists reported elevated numbers of FAs from the SIGO group in particular, suggesting that SIGOs (but not RIGOs) had a negative effect on FAs and only for this sub-sample. Eighth, interview data suggest that RIGOs and SIGOs increased alertness and motivation for a short while, but this did not translate into significant differences in detection rates. Last, the high detection rates for IGOs (Table 15) indicated that participants were generally looking at the display. However, there was no relationship between IGO and TB detection, FAs, number and length of disengagement periods. Therefore the use of IGOs of this nature to ensure that operators are directing their gaze at the display did not significantly contribute to the detection of significant events.

Overall, the IGO intervention was not effective in its current form in terms of enhancing the detection of significant events. This is consistent with the results of Andrew et al. (2003) where significant differences were not found between groups that received a mixed set of IGOs with no relation to significant events and others that represented significant events, versus the control group that did not receive IGOs. Although vigilance performance is influenced by factors such as signal frequency, regularity and probability of occurrence (e.g., Koelega, 1996; Mackie, 1987; Methot & Huitema, 1998), this research demonstrates that performance in CCTV surveillance, and possibly other real life vigilance tasks, is not influenced merely by manipulating the frequency of stimuli requiring a response. However, the IGOs appear to have had a limited positive effect on alertness, motivation and task engagement. In addition, based on the interviews, SIGOs were reported to have a limited effect on attention sets. There are numerous possible reasons for this, based on the nature of complex dynamic displays with scenes from real life, the attention processes involved in detection, IGO characteristics, the manner in which IGOs were implemented, and differences between the TBs used in the research and the significant events.

Both CCTV surveillance and X-ray baggage screening are complex tasks, yet TIP images in baggage screening are more effective than the IGO images used in CCTV in the current research. The IGO intervention was based on TIP for baggage screeners (Neiderman & Fobes, 2005) and ASI for inspection tasks (Wilkinson, 1964). TIP is effective in meeting the objective of evaluating whether screeners are detecting particular threats in bags (Schwaninger et al., 2008). However, results regarding the effectiveness of ASI are mixed (Baker, 1960; Garvey et al., 1959; Wallis & Newton, 1957; Wilkinson, 1964). It is useful to compare the types of tasks, displays and images used in baggage screening and CCTV to facilitate an understanding of the reasons why inserted images are successful in the one context but not in the other.

Both X-ray images and CCTV displays contain complex visual stimuli and embedded objects. However, the images in X-rays are static and in CCTV are dynamic. This is a key difference which could impact on the effectiveness of IGOs. Screeners are able to search each bag systematically because the visual stimuli do not move or change. With CCTV, systematic search of this nature is less likely because the visual stimuli, including people and objects move and change. This is likely to change the operators' priorities regarding the location of their focal attention on a moment-by-moment basis. In addition to this stimulus-driven form of attention control, operators direct their attention on a goal-directed basis. Despite this, with CCTV, systematic searching is likely to be disrupted by dynamics events such as people entering or leaving the scene. The TBs used in the current research were displayed for a maximum of two seconds. Therefore participants were required to focus their attention on continually changing visual stimuli and to catch a moment during their display. This was done without the possibility of reviewing the scene, which is arguably possible with X-rays.

The conditions for inserting images into X-rays and CCTV are also very different. With X-ray, the physical environment is standardised in terms of factors such as lighting and depth of field. Consequently TIP images can be inserted in such a way as to have a realistic appearance and levels of visual conspicuity that are similar to those of real threat objects. With CCTV, this is more difficult to achieve due to factors such as varying depths of field, camera angles and lighting (Gill & Spriggs, 2005). These contribute to the degree of realism and visual conspicuity of the IGOs. In real world operational contexts additional factors such as the operators' ability to operate cameras (Donald, F., 2010) and differing weather conditions (Thiel, 1999) would also affect the extent to which IGOs appear to be realistic. With baggage screening, the X-ray image of each bag is a discrete event, creating the opportunity to project the TIP image onto the X-ray when the X-ray image first appears. This prevents the abrupt onset effect that occurred in the current research with IGOs. However, the continuous video stream, combined with the other factors already mentioned, created challenges in terms of inserting the IGOs in a way that would appear realistic.

The last key difference between inserted images in the X-ray and CCTV contexts concerns the provision of feedback. Screeners using TIP systems receive feedback regarding the detection of the TIP image. When the TIP image is detected, a message appears on the display stating that the image detected was a projected image and asking the screener to examine the bag once again for real threats (Donald et al., 2007). Various studies regarding vigilance (See et al., 1995) and ASI (Mackie et al., 1994) have found feedback to be important in altering vigilance dynamics. Feedback was not provided in the current study because the main purpose was to examine the effect of IGO characteristics on attention processes and the detection of actual significant events. However, the effect of feedback regarding IGO detection and the correct interpretation of SIGOs is recommended for future research.

The design of the particular IGOs used in the current research could have contributed to their lack of influence on TB detection. The current research focused on two types of image characteristics – salience and semantic distance. Salience was chosen for its ability to draw attention (Pomplun et al., 2003) and speed up the detection of the IGOs, thereby reducing the degree to which IGOs distract participants from the rest of the display. Semantic distance was argued to be the most important characteristic in terms of enhancing the detection of TBs. The SIGOs, which had small semantic distances to the TBs and represented the TBs, were intended to assist in the formation and maintenance of relevant attention sets. The RIGOs had random semantic distances and were not related to the TBs. However, the remaining characteristics also required consideration as images consist of combinations of characteristics rather than isolated characteristics.

All RIGOs and SIGOs were visually salient. It was anticipated that highly salient IGOs would be easy to detect and this was considered important in maintaining alertness and reducing potential vigilance decrements. This was based on the 'pop out' effect created by high visual salience (Pomplun et al., 2003; Theeuwes, 1993; Treisman, 1988). Participants confirmed that high salience assisted alertness and concentration due to the 'pop out' effect. Salient IGOs assisted participants by acting as a 'wake-up call' and reminding them of the broad task goals of detecting TBs. Visual salience was increased by the abrupt manner in which IGOs appeared. This is supported by the strong ability of abrupt onsets to capture attention exogenously (Brockmole & Henderson, 2005b; Egeth & Yantis, 1997). However, the subjective increase in alertness did not translate into increased TB detection rates. Potential reasons for this are discussed later in this section.

It is possible that the high salience of IGOs had the potential to distract attention away from important parts of the display with the result that TBs could have been missed. However, this is unlikely in view of the similar detection rates obtained for the benchmark, RIGO and SIGO groups, and the long intervals between IGOs and TBs which allowed participants time to re-focus attention on the video after detecting IGOs. In addition, IGOs were specifically placed so as not to interfere with TB detection. However, the dissociation of IGOs from the TBs could have reduced their potential effects on TB detection. Observation indicated that when participants were focusing on an area far from the IGO, they sometimes missed the IGO, indicating that the IGOs did not distract them when their attention was focused elsewhere. This is consistent with the contingent involuntary orienting hypothesis whereby goal-directed search overrides stimulus-driven attention capture (Folk et al., 1992).

Differences in salience could partially account for the lack of relationship between IGO and TB detection rates. Some participants reported a high number of IGOs but very few TBs and are likely to have adopted a search process similar to the singleton detection mode (Duncan, 1984). It is possible that these participants focused almost exclusively on the search for salient IGOs to the detriment of TBs. This suggests that a set of IGOs with varying salience and detection difficulty levels would be preferable to using IGOs with consistently high or low salience levels. At a practical level, the lack of relationship between IGO and TB detection implies that IGO detection rates cannot be used to predict TB detection in a fair, valid or reliable manner, and should therefore not be used in performance management systems for CCTV. In addition, it is meaningless to use salient IGOs to evaluate whether operators are looking at the display because the detection of salient IGOs could mask the lack of active search and visual analysis required to detect significant events.

Semantic distance was the key IGO characteristic aimed at enhancing detection rates. Small semantic distances between the content of SIGOs and TBs in the video were intended to increase TB detection rates through attention sets and endogenous attention guidance. This is based on the importance of knowledge of what one is searching for in assisting vigilance performance (Davies & Parasuraman, 1982) and visual search (Castelhano & Heaven, 2010; Gilchrist & Harvey, 2000; Pomplun et al., 2003). Knowledge of significant events assists endogenous attention control via the attention set (Leber & Egeth, 2006; Most et al., 2005; Wickens & McCarley, 2008). Therefore the attention set is crucial to successful search (Most et al., 2005). Based on this, it was envisaged that SIGOs would enhance the detection process by depicting TBs, creating knowledge of what TBs look like, and influencing attention sets. As RIGOs were only intended to maintain alertness, they had random semantic distances and their content was not related to TBs.

Interviews supported the importance of semantic distance as an IGO characteristic, with different qualitative responses being obtained for RIGOs and SIGOs in this regard. Differences were noted in terms of emotive responses and comments on attention guidance. Some participants were annoyed by RIGOs because their content was not related to the detection task or what they were looking for. These participants perceived the random semantic distances of RIGOs as being intrusive and distracting. No such comments were made for SIGOs, where the short semantic distances were viewed positively. Participants reported that the small semantic distances of SIGOs had a positive influence on the search process by guiding them in terms of what to look for. This suggests that SIGOs influenced attention sets and is consistent with the importance of attention sets in visual search (Most et al., 2005). However, the reported benefits of high salience and small semantic distances did not translate into improved TB detection rates. Several possible explanations are offered. An appropriate attention set is only one of many factors involved in detection and other processes such as broader SA (Endsley, 1995), appropriate eye movements (Castelhano & Henderson, 2008; Henderson et al., 2007; Foulsham & Underwood, 2008), attention deployment (Endsley, 2000), visual analysis (Donald & Donald, 2008), and recognition (Cole & Liversedge, 2006; Gibson & Peterson, 2001) also contribute to successful detection.

The lack of impact by the IGOs could have been related to insufficient intensity and duration of their influence on attention sets. This is likely in view of the dynamic nature of the display and the need to update the contents of working memory on a continual basis in order to maintain SA (Endsley, 2000). In addition, SIGOs could have created attention sets that were too narrow to assist in the detection of TBs. It is possible that narrowly defined attention sets primed participants for specific TBs to the exclusion of TBs which took place in other forms (e.g., different postures). When TBs with different features from those represented in the last SIGO appeared, participants may have missed them because they were not expecting them. This is consistent with inattentional (Mack & Rock, 1998) and change blindness studies (Rensink, 2002; Simons & Levin, 1997). In this situation, participants would have expected SIGOs to act as accurate cues for TBs. The high level of FAs for generalists in the SIGO groups is consistent with this. In addition, improved attention sets may have been undermined by the depletion of attention resources and difficulty in re-engaging with the task. This is consistent with multiple resource theories of attention (Wickens & McCarley, 2008).

The responses required when SIGOs were detected could have influenced the degree of visual analysis that occurred and the subsequent influence on the attention set. The SIGOs depicted TBs being performed in various ways, such as the side view of a person crouching down or a rear view of a person bending from the waist and picking up a small object. The use of different postures was intended to signify different ways in which TBs could occur. Participants were required to name the relevant TB. However, it may have been more effective if participants had been asked to describe the posture in addition to naming the

TB. This would have required additional visual analysis and may have contributed to the formation of more accurate attention sets and possibly mental models. It may have provided a greater understanding of what to look for and assisted in bridging the gap between the detection of SIGOs and actual TBs.

An explanation for the failure of IGOs to influence the detection of TBs that is related to the nature of attention sets, is that IGOs had insufficient influence on SA. Based on Endsley's (2000) levels of SA (perception, comprehension and prediction of events), interview data suggest tentatively that SIGOs assisted some participants in developing a low level of SA which assists people in perceiving important events. However, this is unlikely to be sufficient in the CCTV context where significant events can occur in a split second and involve small, subtle movements. To detect these events, the eyes need to be directed at the relevant location when the events occur. This implies that operators need to anticipate and predict significant events so that they are looking at the right place at the right time. This is supported by the finding that operators with high levels of SA are more likely to perceive, recognise, and comprehend significant events than those with low levels of SA (Federico, 1995).

Higher levels of SA are required for the prediction, recognition and comprehension of significant events (Federico, 1995). If IGOs only assisted in the development of lower levels of SA, they would not have assisted in the prediction and recognition of TBs. Based on this argument, SIGOs need to contribute to the development of high levels of SA in order to assist with the detection of TBs. In addition to developing an attention set for TBs, they would need to assist in creating an understanding of suspicious events that often precede and predict significant events. This is consistent with the critical role of accurate SA in making correct decisions (Wright et al., 2004). However, SA develops from a variety of information sources (Endsley, 2000) and the current research demonstrates that SIGOs representing small semantic distances to TBs alone are insufficient in developing adequate SA. The knowledge involved in SA is embodied in mental models which guide attention to relevant events (Nunes & Mogford, 2003). It is not known whether IGOs assist in developing mental models of significant events and this requires future research.

It is possible that a set of SIGOs representing suspicious events as well as TBs would assist the detection of TBs and more fully fledged significant events more effectively than SIGOs that only focus on narrow TBs. (Effective training courses could probably do this as easily, and the use of training followed by IGOs on the job could possibly reinforce the benefits of training). An understanding of behaviours assists in anticipating events and recognising them as they unfold (Donald, 2004). Images of suspicious behaviour are likely to contribute to the comprehension of behaviours and patterns of events that predict significant events as they would provide observers with additional information. This could assist in developing broader and more accurate attention sets and higher levels of SA. SIGOs of this nature are more likely to assist in developing a broader frame of reference for the searching process, rather than focusing only on a narrowly defined attention set and specific TBs. This is supported by the finding that specialists in particular used suspicious events to cue and predict TBs.

The static nature of SIGOs is likely to have impacted on the semantic distance and concreteness of the IGOs. The semantic distances of the SIGOs may not have been small enough to represent the TBs in a realistic manner because SIGOs did not contain moving parts that replicated the dynamic properties of TBs. The static nature of SIGOs reduced their realism and concrete depiction of movements. The use of dynamic SIGOs that moved in a realistic manner (e.g., SIGOs based on filmed footage) may have been more effective than static SIGOs as their similarity to real TBs would have been higher. This might have created smaller semantic distances between the SIGOs and TBs, and would have increased the concreteness and realism of SIGOs. SIGOs that represent movement are different from safety and warning signs that incorporate dynamic properties such as flashing lights. With warning signs, dynamic properties and other salient features are aimed at drawing attention to the signs (Wogalter et al., 1987) rather than depicting how target movements occur, as intended with SIGOs.

Dynamic SIGOs containing realistic movement are not without potential disadvantages. They are likely to be missed or misinterpreted, especially if the significant events they depict are played at realistic speeds. This is because many significant events occur quickly, are subtle and are difficult to detect as evidenced by the low detection rates in the current research. If SIGOs are misunderstood or missed, they are unlikely to facilitate the detection of significant events and this would undermine the intervention. Therefore ways of representing significant events in a dynamic but effective manner need to be identified. Slow motion is one such option, but reduces the realism of SIGOs. Alternatively, SIGOs could initially be introduced in slow motion and then speeded up to realistic speeds over a period of time. The second potential disadvantage of IGOs with moving parts is that they are likely to increase distraction from significant events and fatigue. This is based on the tendency of certain types of motion to draw attention (Chastain et al., 2002; Franconeri & Simons, 2003) and the depletion of attention resources described by multiple cognitive resource theory (Wickens, 2002).

The interaction between semantic distance, concreteness and realism illustrates how the characteristics combine in images and cannot be examined separately. This is consistent with complex vigilance tasks where multiple facets are inter-related and cannot be studied in isolation (Howell et al., 1966). Although the research focused on semantic distance and salience, other characteristics were also included in the IGOs as they are integral aspects of images and cannot be excluded. Concreteness, complexity, novelty and realism were considered in IGO design. All IGOs were concrete and none contained abstract features. This may have influenced the ability of SIGOs to influence TB detection. Research indicates that concrete images are more readily understood than abstract images and their comprehension is less reliant on learning (Familant & Detweiler, 1993; McDougall et al., 2001; McDougall & Isherwood, 2009). Therefore it seemed logical to use concrete images, especially as visual literacy levels were likely to vary amongst participants. However, SIGOs were static representations of dynamic TBs, and only captured one moment in time of behaviours that consisted of sequences of movements. It is possible that the addition of abstract features, such as arrows, to the concrete images would enhance the accurate depiction of movement. Alternatively, the use of SIGOs with moving parts could possibly have addressed this issue, as discussed above.

Complexity, novelty and realism are likely to have influenced the effectiveness of IGOs in enhancing TB detection. SIGOs were more complex than RIGOs as they were based on photographs and involved more parts, colours and lines. It is possible that simpler SIGOs would have had different effects on attention processes and comprehension. This is based on findings that simple images tend to be more salient (Wogalter et al., 1987) and recommendations that the images used in safety signs should be simple (Gittins, 1986). However, the advantages of simplicity need to be weighed against the need for comprehension of the SIGOs and the context in which they are placed. SIGOs are inserted into complex visual scenes where visual stimuli (e.g., people, objects) have particular physical features (e.g., the types of clothing worn and equipment used). The SIGOs were designed so that their visual features were congruent with those encountered in the display. This was intended to increase realism, assist in the comprehension of the SIGOs, and to facilitate the recognition of actual TBs in the display.

In addition to the use of static IGOs, realism was reduced through the placement of IGOs in locations that were incongruent with the meaning of display. Some research has demonstrated that objects that are incongruent with the gist of scenes are detected quickly (Davenport & Potter, 2004; Gordon, 2004). This suggests that IGOs placed in incongruent locations would be detected quickly and would reduce the duration for which IGOs distract attention from the rest of the display. However, other studies have found that incongruent stimuli are not detected efficiently (De Graef et al., 1990; Henderson et al., 1999). Incongruent objects have been shown to be detected more quickly than congruent objects only when they are also visually salient (Foulsham & Underwood, 2008). As the IGOs were salient, they are likely to have drawn attention quickly. However, the incongruent placement could have reduced comprehension for participants with low levels of visual literacy. This is supported by limitations in people’s understanding of images (Dowse & Ehlers, 2004; Hancock et al., 2004; Houts et al., 2006; Mansoor & Dowse, 2007; Miller & Stanney, 1997; Wilkinson et al., 1997). The response format in the current research required participants to state which TBs were depicted in SIGOs, ensuring that participants understood which TB was depicted. However, for some participants, the incongruent location could have decreased their understanding of how the TB depicted could occur.

The more realistic placement of SIGOs in scenes would have allowed them to ‘blend’ into the scene to a greater extent, making them less salient and more difficult to detect. This could have had the advantage of requiring participants to engage in a more active searching process to detect them, but would also have increased mental workload. The mindlessness theory of vigilance predicts that the increased challenge would have enhanced performance (Robertson et al., 1997), but multiple resource theories of attention predict that attention resource depletion would occur in these circumstances (Smit et al., 2004; Wickens, 2008). Further research on the effects of the location of IGOs and their congruence with the scene's meaning is needed.

The insertion of IGOs in a completely realistic manner would need to take various additional factors into account, such as depth of field, perspective, lighting, and size of IGOs. These influence the degree to which IGOs blend with the gist and physical properties of the scene (salience), and whether they are likely to obstruct significant events that occur at the same time. Future research is required to identify the effects of these factors on the ability of IGOs to enhance the detection of significant events.

Much of the above discussion on dynamic properties and the realism of IGOs is relevant to the way in which IGOs are inserted into displays, or the implementation strategy. In addition to designing IGOs based on appropriate characteristics, the way in which IGOs are inserted into displays is likely to influence IGO effectiveness in facilitating the detection of significant events, search efficiency, attention resources required, fatigue, and potential distraction. Dynamic and spatiotemporal factors are key aspects of such an implementation strategy. Spatiotemporal factors are important because they affect the predictability of targets and expectancies regarding occurrence, both of which influence detection (Mackie, 1987). Based on SDT, expectancies alter the response criterion which, in turn, affects detection rates (MacMillan & Creelman, 1991), as explained in chapter 2. In order to manage expectancies, IGOs were inserted irregularly and in different locations (although RIGOs and SIGOs were inserted at the same times and locations as each other). IGOs were mostly placed in locations that did not interfere with activities being shown in the video. It is important to note that IGOs did not function as spatial or temporal cues as they did not signal where or when TBs would occur. This is in keeping with real life CCTV systems where it would not be possible to predict where and when significant events occur. SIGOs only indicated the types of significant events that could occur at unknown times and locations. Therefore they did not reduce spatiotemporal uncertainty and this could have reduced their effectiveness. However, in real world contexts significant events are not known and it is therefore not possible to insert the 'right' IGO before a significant event.

Spatiotemporal uncertainty is related to decreased vigilance performance (Koelega, 1996) and increased workload (Parasuraman & Mouloua, 1996) and scanning (Liu, 1996). This suggests that IGOs should be inserted at predictable locations and frequencies. However, the potential benefits of predictability were weighed against potential disadvantages. An insertion schedule with predictable spatiotemporal factors is likely to create accurate expectancies regarding where and when IGOs are found. As expectancies influence behaviour (Most et al., 2005), operators are likely to adjust their search patterns and look at the appropriate location when they expect IGOs to occur. This could result in a high hit rate for IGOs, but would not necessarily encourage active search of the rest of the display at other times. Therefore significant events may be missed. This is particularly likely if operators are rewarded for IGO detection or if the perceived consequences of missing significant events are not seen as being serious. At an intuitive level it seems appropriate that IGOs are inserted at irregular times and locations.

In summary, IGOs did not achieve the intended results. Reasons for this include the complex and dynamic nature of the displays and detection task, the ability of IGOs to influence attention sets, SA and attention guidance, the depletion of attention resources, the way in which IGOs were inserted into the display and their static nature, and the need for feedback. Highly salient IGOs were effective in drawing attention to themselves despite being inserted into complex backgrounds. However, salient IGOs only had a limited effect on alertness and did not influence the detection of significant events. Sets of IGOs with mixed salience levels are likely to be more effective in encouraging active searching processes.

Based on the interviews, it appears that semantic distance is an important characteristic for IGOs. SIGOs had a limited effect on attention sets and assisting participants in knowing what to look for. It is unknown whether SIGOs contributed to the development of SA, mental models of TBs, or scan patterns, as the research did not set out to evaluate these directly. It is concluded that SIGOs did not assist visual analysis sufficiently. IGOs with small semantic distances to significant events have greater potential to influence the detection of significant events than IGOs that represent content that is not related to the detection task. However, it is recommended that the content of IGOs be broadened to include suspicious events that predict significant events. This is aimed at contributing to the development of higher levels of SA that can be used to comprehend what is occurring in displays and predict significant events. Given the complexity of vigilance and detection processes, it is unlikely that a single intervention will address all issues (Donald & Donald, 2008). This is consistent with the fact that vigilance consists of multiple processes (Koelega, 1996) and therefore numerous processes need to be addressed for interventions to be effective. It is likely that IGO systems need to supplement other interventions, such as improved selection and training.

More direct interventions aimed at enhancing visual analysis, SA, active searching, scan patterns, the self-regulation of attention and task engagement are more likely to have an impact on detection rates. However, IGOs could possibly be used to reinforce and maintain the effects of these interventions. In addition, the detection rates of operators need to be considered in terms of what is realistic and how job design and technology can better support operators (Dubbeld, 2005).

In terms of the IGO intervention, the current research has contributed to theory by identifying and categorising a range of characteristics for consideration in IGO design (chapter 5) and proposing ways in which IGOs could influence attention processes (chapter 4). It confirms the role of salience in assisting the detection of IGOs themselves. The emphasis on the importance of semantic distance to significant events represented in IGOs is a key contribution. No previous research in the CCTV context has examined the role of semantic distance and certain researchers (e.g., Neil et al., 2007) have only used random images with no relationship to the detection task.

The research highlights the importance of operators' frames of reference for the detection task. This goes beyond a narrowly defined attention set. Based on Endsley's (2000) levels of SA, a broader frame of reference includes higher levels of SA that can be used not only to perceive potential significant events, but to comprehend what is being observed and predict significant events. SA in the context of CCTV has previously been neglected by researchers and needs to receive more emphasis. The role of active search is a unique contribution made by the research and deserves further research. Although visual analysis has been referred to in previous studies (Donald & Donald, 2008), the processes involved require more research attention.

It is evident that CCTV surveillance is a complex and demanding task and that significant events are likely to be missed, regardless of whether a vigilance decrement occurs or not. This suggests that researchers' emphasis on the vigilance decrement to the detriment of overall vigilance levels in the past has been somewhat inappropriate for complex, real world vigilance tasks. The overall level of vigilance is as important as the vigilance decrement in the CCTV context. This supports a similar conclusion reached by Davies and Parasuraman (1982) that the overall vigilance level has been neglected. Existing literature regarding the vigilance decrement has made a major contribution to the area by identifying broad processes that are required to sustain attention. However, it has been criticised for saying little about what takes place during the vigilance task (Koelega, 1996). The current research assists in bridging this gap by highlighting the important role of active search, visual analysis, fluctuations in task engagement, and motivation in surveillance and detection. It suggests that interventions that address a number of these processes would be beneficial.

Previous researchers have treated CCTV surveillance as a unitary task. However, this research found differences in work exposure and experience that contributed to detection rates, performance fluctuations and the ability to maintain task engagement over time. The identification of different surveillance tasks implies that both CCTV and other vigilance research tasks and samples need to be carefully delineated in order to ensure ecological validity. This is a major departure from previous CCTV research which tends to group all CCTV surveillance tasks together.

The next section explains the limitations of the research.

8.4. Limitations of the Research

The discussion of limitations is important in any research as it highlights aspects that could limit the validity of the findings and point to alternative approaches (Rosnow & Rosenthal, 1996). Limitations of the current research are discussed in terms of an alternative research design, the nature of the sample, participants' expectations, measurement issues, and differences between this research and operational CCTV surveillance contexts.

The research design informs the ability of the study to answer research questions (Rosnow & Rosenthal, 1996). An alternative design, in the form of a cross-lagged or panel design was considered. This consists of each participant undergoing each treatment and therefore acting as his/her own control (Rosnow & Rosenthal, 1996). For the current study, each participant would have observed thirty minutes each of the benchmark, RIGO and SIGO conditions, with participants viewing these in different sequences. The use of SAMAE scores would not have been necessary in allocating participants to groups, representing a significant time saving. However, a cross-lagged design was not considered appropriate because of the potential for the different treatments to contaminate each other and for a vigilance decrement to influence performance between time phases.

It is possible that the length of the vigilance task influenced the results. The ninety-minute period was longer than the period typically used in vigilance studies which is around thirty (Molloy & Parasuraman, 1996; Parasuraman & Mouloua, 1987) to forty minutes (Colquhoun & Edwards, 1970). Many studies have used shorter tasks (Matthews, Davies, & Holley, 1993; Temple et al., 2000). However, the Andrew et al. (2003) study used a task that lasted over three hours and found that performance decreased only after two hours. It is possible that a more pervasive vigilance decrement would have occurred in a longer task.

There were several limitations in terms of the sample. Compared with studies using self-report questionnaires and survey, the sample size appears limited. However, small samples are typically used in time intensive research. Vigilance research typically uses small samples, such as 12 (Caggiano & Parasuraman, 2004) to 30 (Parasuraman & Davies, 1976), although a review of the literature indicates that 30 is on the large side. Visual search research typically uses samples of 5 to 15 participants (Wolfe, 1998). Research on task engagement in the context of sustained attention also uses small samples (e.g., 21 in Smallwood et al., 2004), but larger samples have been used when shorter tasks (Cheyne et al., 2009) or other sources of data are used, such as archival data and self-report questionnaires (Carrière et al., 2008). The current sample of 73 is therefore considered to be a large sample in view of the task length and disruption from work activities. CCTV control rooms tend to be fairly small and specialised operations. This tends to restrict the available sample from any one operation, hence the need to draw participants from a number of organisations in the current study.

Despite the relative largeness of the sample in view of the time demands on participants, the sample size reflects difficulties in obtaining larger samples in real life settings (Pigeau et al., 1995). CCTV control rooms require relatively small numbers of operators (Donald, 2001), and participation in the research disrupted standard work schedules, shifts and relaxation time. Further, it could have resulted in loss of product through lack of operational performance due to employees being removed from their workstations to participate in the research. This limited the number of operators who participated. Research in real life environments tends to be constrained by practical demands. While this gives rise to limitations in the research, the advantages of the application of research in real world settings remains important (Koelega, 1996).

The exclusion of a large proportion of operators from analyses of the results is problematic. These participants were excluded on the basis of very high levels of FAs on SAMAE or the video which were seen as rendering their results meaningless. This is of particular concern as the intervention is aimed at operators. Possible explanations for the high levels of FAs were provided in section 6.1.2. The exclusion of these operators suggests that selection, training and visual analysis skills affect performance in a number of settings and could lead to performance enhancements if addressed.

Issues related to language could have affected the results of the study. The research was conducted in English which was supplemented with Afrikaans when necessary. However, the first language of many participants was another of the eleven official languages of South Africa. This meant that extra care was taken to ensure that all participants understood what was required of them. Additional explanations were provided when necessary. This undermined the standardisation of instructions on video, but comprehension was considered more important than standardisation. English and Afrikaans were used by all the organisations to conduct business and were therefore understood to an adequate degree by all participants.

Participants' expectations could have influenced the results particularly with regard to the end spurt effect. However, from an ethical and practical perspective it was inappropriate to withhold knowledge of the task duration. A lengthy period of time was being requested of participants who had other commitments and who were at times being taken off their jobs to take part in the research.

The process used for allocating difficulty levels to TBs is a limitation of the research. The research required the use of TBs that occurred in different ways, as well as varying detection difficulty levels as these are inherent to CCTV surveillance. This differs from laboratory vigilance research where there is only one or a few targets and the detection difficulty level remains constant (e.g., Broadbent, 1953; Mackworth, 1957, 1970). This also applies to much visual search and attention capture research which uses simple stimuli with consistent detection difficulty levels (e.g., Theeuwes, 1993; Treisman, 1988; Treisman & Gelade, 1980; Wolfe, 1998; Yantis, 1996). The use of a detection difficulty level introduced a unique variable to the research. The difficulty level was based on ratings and actual detection rates from the pilot studies. However, it is recommended that the calculation of difficulty levels should allow for finer distinctions in future research.

The small number of different types of TBs for detection in the current study could have impacted on the effectiveness of the IGOs. The use of only four different types of TBs is likely to have reduced the need for stimuli that 'prompt' working memory. This is supported by findings on working memory indicating that the optimal number of chunks to maintain in working memory simultaneously is two to seven (Edin, Klingberg, Johansson, McNab, Tegnér, & Compte, 2009). These chunks could include groups of objects, such as the different postures used for the different TBs in the current research. If a significantly larger set of TBs had been used, SIGOs may have had more opportunity to activate relevant schemas and assist in maintaining them in working memory. Additional TBs were considered, but were excluded because they were highly ambiguous, required high levels of SA for detection, or were extremely difficult to detect and therefore unlikely to be reported by participants from different industries.

An important measurement issue concerns accurate measurement of attention processes, including disengagement. These can be evaluated and inferred from measures such as eye movements (Foulsham & Underwood, 2007, 2008) cerebral blood flow (Deutsch, Papanicolau, Bourbon, & Eisenberg, 1987; Hitchcock et al., 2003), electroencephalogram (Lal & Craig, 2002, Wright & McGown, 2001), galvanic skin response (Loeb & Alluisi, 1984), positron emission tomography (Pardo, Fox, & Raichle, 1991), and various other biological indices. Some of these are intrusive (Ibarra-Orozco, Gonzalez-Mendoza, Hernandez-Gress, Diederichs, & Kortelainen, 2008) and were likely to reduce the number of people who were willing to participate. Alternatively, more qualitatively intensive methods such as concurrent verbal protocols (Ericsson & Simons, 1993) (which would also have interfered with task performance), in-depth interviews (Kaar, 2007) or retrospective protocols (Rijlaarsdam, van den Bergh, & Couzijn, 1996) supplemented with video-recordings could have been used. Despite the limitations of these approaches, they could have contributed significantly to an understanding of the CCTV viewing and detection process and task engagement.

In the absence of such measures, periods of disengagement were inferred from observations of non-verbal behaviours in combination with periods when nothing was reported. Consequently the evaluation of disengagement periods is a crude measure and results should be treated with caution. In addition, superficial disengagement that was not accompanied by non-verbal behaviour was not identified. Therefore disengagement was probably under-estimated. Despite this, the identification of disengagement periods pointed to the need for future and more accurate research on task engagement and renewal of interest in surveillance contexts. Physical measurement of attention processes (Smallwood et al., 2004) may be considered to assist in establishing disengagement.

The last set of limitations concern differences between this research and real world CCTV surveillance contexts. These are important as they inform the ability of the research to generalise to real world situations. The most striking difference is that participants only observed one display in the research, but in operational contexts are typically responsible for three to thirty cameras simultaneously, often on multiplex displays. Larger numbers of cameras are associated with lower detection rates (Neil et al., 2007; van Voorthuijsen et al., 2005; Wallace & Diffley, 1998; Wallace et al., n.d.). Therefore observing one camera should have required fewer cognitive demands and should have improved detection rates. The intervention could have different effects when multiple displays are used, especially if they draw attention to displays that tend to be neglected in the searching process.

The above limitations represent some of the difficulties in conducting research on real world tasks compared with experiments using simple, synthetic stimuli. These challenges may partially explain why researchers have neglected CCTV surveillance as an area of research. Despite this, CCTV provides a rich context for research into vigilance, visual search, attention processes, and task engagement.

8.5. Directions for Future Research

A number of recommendations for future research follow, focusing on the regulation of attention and engagement, fluctuations in vigilance over time, IGO interventions, the overall level of vigilance performance, the viewing and detection process (including visual analysis and active search, strategies for searching), and types of CCTV surveillance.

The regulation of task engagement emerged as a process that requires future research in the context of complex, dynamic and sustained tasks. A greater research focus on individual fluctuations in performance over time rather than the traditional vigilance decrement curve would create an understanding of processes that reduce overall detection rates, as well as the detection process and mechanisms of task disengagement and re-engagement. Factors that moderate task engagement at both the task and individual levels need to be identified. Certain factors have been identified with simple tasks (Northern, 2010; Smallwood et al., 2004) but their applicability to complex tasks requires investigation. A greater understanding of the antecedents of task disengagement and re-engagement would contribute to designing jobs and developing appropriate interventions. The use of biological and neurological measures (Pope, Bogart, & Bartolome, 1995; Smallwood et al., 2004) of task engagement is likely to add value to research in this area. Although disengagement probably cannot be completely prevented in lengthy tasks, the challenge is to understand and intervene in the processes of re-engagement, renewing interest and effort. Knowledge of the mechanisms for re-engagement could form the basis for an alternative intervention aimed at improving vigilance performance.

The IGO intervention in its current form is limited in its ability to significantly influence detection rates. However, with further development it may have the potential to achieve this. It is recommended that future research focus on the effects of feedback regarding IGO detection, ways of designing IGOs so that they have a stronger and longer lasting effect on alertness, contributing to improved visual analysis, perception and recognition, and developing a broader frame of reference for the detection task. A number of recommendations follow regarding IGO characteristics and the way in which they are implemented.

Research on ASI (Baker, 1961; Wilkinson, 1964) and vigilance (Mackie et al., 1994) demonstrates the importance of feedback in vigilance performance. In addition, feedback is incorporated into TIP systems (Donald et al., 2007). Feedback on the detection of IGOs and related rewards and sanctions are likely to influence the detection of IGOs and significant events. Where only IGO detection (and not the detection of significant events) is evaluated, operators may focus on IGOs and neglect significant events. Therefore, future research should consider feedback and broader performance management strategies as part of the implementation of potential IGO systems. Research into the influence of feedback regarding IGO detection on significant event detection levels, performance fluctuations over time, task engagement and motivation could provide insight into attention processes and improve IGO interventions.

The current research focused on salience and semantic distance. Salience made a positive contribution in terms of alertness. However, high salience could have reduced the need to search the display actively, or to ‘look harder.’ Therefore it is recommended that salience be manipulated to examine the effects of different levels of salience on alertness and the detection of significant events in complex displays. It is probable that IGOs with lower salience would encourage more active search. However, they would also have the disadvantage of requiring additional attention resources and effort, thereby increasing fatigue and possibly task disengagement. These possibilities require research. An additional option is to experiment with a set of IGOs with varying salience levels. This could retain the potential benefits of increased alertness and motivation for highly salient IGOs and active search required by low salience IGOs.

The concept of semantic distance has value for IGOs but could be broadened in future research. This would enable IGOs to represent suspicious events associated with significant events which operators could use as cues for significant events. This would allow for the development of a broader frame of reference and attention sets that are not narrowly defined. This is particularly relevant to dynamic significant events that are highly transient. The contribution of such IGOs to the development of various levels of SA needs to be examined. Related to this is the need for research on IGOs using a large set of TBs or significant events where a broader frame of reference is required for detection than was the case in the current research.

Further research into image characteristics in addition to salience and semantic distance (complexity, concreteness, novelty and realism) would assist in developing the intervention. The current research used concrete images. However, the contribution of abstract features (e.g., arrows) needs to be evaluated. These could assist the comprehension of static images that depict movement and highlight the types of movements to search for, thereby facilitating the formation of accurate attention sets. This needs to be examined in terms of visual literacy levels. This is based on the finding that the meaning of images is learnt (McDougall & Isherwood, 2009) and that image comprehension is related to culture and is not universal (Dowse & Ehlers, 2005; Rother, 2008). This is particularly important in view of South Africa's diversity of cultures.

The current research included specific spatio-temporal and dynamic properties for the IGOs. Different ways of inserting IGOs into displays could potentially have different effects on attention processes and require research. These include the frequency and regularity of IGO insertion, location, the manner in which they first appear (i.e., abrupt onset versus more realistic ways of appearing), and other dynamic properties (e.g., moving parts, different ways of moving across the display). The use of animated, enacted and real footage should be evaluated. IGO characteristics and their implementation strategy influence salience and congruency with the scene's meaning. These in turn are likely to affect realism. However, the design and implementation of highly realistic IGOs is challenging as it requires consideration of all factors discussed as well as camera angle, lighting, size and depth of field, amongst others.

In the current research, all IGOs were static. However, IGOs could include a variety of dynamic properties that could influence attention and search processes. Three broad categories of movement require consideration - how IGOs first make their appearance in the display, movement of IGOs in the display, and movement of parts of the IGO itself. This is similar to the inclusion of dynamic properties in warning signs (Laughery, 2006). The first two categories are more likely to influence how quickly IGOs are detected than how they could facilitate the detection of significant events. This is based on the finding that abrupt onsets exert strong exogenous attention capture (Brockmole & Henderson, 2005b; Egeth & Yantis, 1997) and that certain movements tend to draw attention (Chastain et al., 2002; Franconeri & Simons, 2003). However, future research is needed to investigate the effects of IGOs' dynamic properties on the detection of significant events and potential distraction from real significant events.

The insertion schedule in the current research considered factors such as IGO frequency in relation to the duration of the task. IGOs were displayed for four seconds, but it is possible that different lengths of IGO display could have influenced detection processes differently. The effects of different insertion schedules for IGOs need to be examined. This is based on the association between vigilance performance, signal predictability, event rates and the frequency of change in displays (Ballard, 1996; Wickens, 1992). Factors to consider in an insertion schedule include the frequency and regularity of insertion, duration of IGOs on the display, intervals between IGOs and frequency patterns. When IGOs are inserted very frequently, they are likely to demand attention resources that would otherwise be devoted to significant events. This is based on the multiple attention resource theory (Wickens, 2000). However, the mindlessness theory of vigilance (Robertson et al., 1997) predicts that greater challenge and mental workload encourage active engagement in vigilance tasks (Forster & Lavie, 2009). In addition, frequent insertion could contribute to the perception that IGO detection is more important than the detection for significant events, and subsequent neglect of searches for significant events. Infrequent insertion of IGOs, on the other hand, may be insufficient in contributing to an appropriate frame of reference. The effects of different insertion schedules are currently unknown and require investigation.

An alternative approach to the timing of IGO insertion is to examine the effects of flexible insertion schedules. Two such approaches are proposed. The first involves inserting IGOs at times when they are most likely to be beneficial, such as during periods of low activity and in less cluttered scenes. Second, insertion could be linked to biocybernetic feedback when there are indications that alertness has decreased. Biocybernetic feedback has been used effectively to monitor alertness in other contexts (Pope et al., 1995). Schedules tailored to individual alertness and disengagement periods are likely to be more effective than standard schedules that are applied despite individual variations in task engagement. This is supported by the TIP frequency schedule which is based on workload in the form of the number of bags being processed rather than the passage of time (Cutler & Paddock, 2009). The frequency of TIP insertion varies according to circumstances (Neiderman & Fobes, 2005).

The recommendation that research be conducted on a number of processes involved in complex vigilance tasks suggests that no single intervention will address all processes and issues related to detection accuracy. In addition to further research on IGOs, alternative interventions could assist. In particular, research is recommended into interventions aimed at improving visual analysis skills, increasing SA and the ability to recognise, comprehend and predict significant events (Endsley, 2000), encouraging active search and interrogation of displays, developing effective search strategies (Di Stasi et al., 2011), and regulating task engagement (MacLean et al., 2009; McVay et al., 2009).

There is currently little understanding of how the many cognitive processes involved in complex and dynamic vigilance tasks interact in the searching and detection processes. The roles of broader frames of reference, specific attention sets (Most et al., 2005), SA (Adams, Tenney, & Pew, 1991; Endsley, 2001), visual analysis, search strategies (Wells et al., 2006), eye movements (Di Stasi et al., 2011; Hauland, 2008; Henderson, 2003), task engagement (Northern, 2010; Smallwood, et al., 2004), and active search need to be better understood for real world vigilance tasks. With a few exceptions (e.g., research on visual search using photographs of natural scenes (Brockmole & Henderson, 2005a) and on certain complex tasks such as air traffic control (Nunes & Mogford, 2003), little is known about these processes in the context of complex and dynamic tasks. The concept of SA needs to be parsed into its constituent processes in the context of CCTV surveillance so that there is a better understanding of the viewing and detection process. This would assist in addressing some of the criticisms regarding the concept of SA made by researchers such as Flach (1995) and Dekker and Hollnagel (2004).

With CCTV and other complex and dynamic vigilance tasks, little is known about how operators prioritise areas for search and decisions regarding where to direct their gaze. Studies using eye movement technology have been useful in other contexts (Di Stasi et al., 2011; Hauland, 2008; Henderson, 2003) and are recommended for surveillance tasks. Combined with expertise comparisons (Ericsson, et al., 2006; Jarodzka et al., 2010; Walker et al., 2010), these are likely to yield greater understanding of intentional searching and attention processes. From an operational perspective, it would be useful to understand which patterns of eye movements are most effective and how these differ between experts and novices.

The identification of active and passive searching processes in the current research prompts the need to research into the interrelationships between effort, motivation, working memory capacity, attention resources, SA, and task engagement. Although early vigilance studies examined some of these variables (Kahneman, 1973), theories need to be applied to complex tasks. More insight is needed into the search strategies involved in active search. This includes the interactions between endogenous and exogenous attention (Arnott & Pratt, 2002; Olivers, Spalek, Kawahara, Di Lollo, Dux, Asplund, & Marois, 2009), as evident in the roles of attention sets (Most et al., 2005) and SA (Adams et al., 1991; Endsley, 2001), and how these interact with the physical properties and content of scenes. Further knowledge of the distinctions in searching strategies between experts and novices in this regard would be useful.

Visual analysis is a key skill in the detection process, yet the processes involved in it are poorly understood. Although visual analysis has been applied in a number of contexts (e.g., Lowe, Hoffman, DeLong, Patz, & Coleman, 1994), research focuses on the outputs of the visual analysis rather than the perceptual and cognitive processes that combine in the process of visual analysis. Existing theories of visual perception (e.g., Bundesen, 1990; Itti & Koch, 2000; Treisman & Gelade, 1980; Treisman & Sato, 1990; Wolfe, 1994) assist in explaining how perception occurs at a sensory level, but provide little information regarding intentional visual analysis and search processes. The components and factors involved in visually analysing complex scenes require further understanding.

It has long been established that vigilance tasks differ in their demands and types (Warm et al., 2008). However, researchers have tended to treat the CCTV detection process as a unitary task despite recognition of the different demands made by different CCTV technology and viewing conditions (Müller & Boos, 2004; Tullio, et al., 2010). The different types of work exposure provided by various CCTV contexts and their associated variation in detection complexity suggest that an empirically validated typology of CCTV surveillance tasks needs to be developed. This would contribute to an understanding of differences in the detection process between contexts and form the basis for comparison of research studies, similarly to typologies of vigilance tasks (e.g., Donald, F, 2008; Parasuraman & Davies, 1977). In addition to factors that influence complexity (Donald, F., 2010), it could consider aspects such as proactive, real-time monitoring versus reactions to alarms, reviewing of footage, detection difficulty levels, and the nature of significant events that occur. This would enable specification of the context and nature of CCTV surveillance, increasing the ecological validity of research.

CCTV surveillance operators frequently perform additional job-related tasks while monitoring displays. This requires the switching of attention between tasks. The effects of task switching (Wickens & McCarley, 2008) on the viewing and detection process need to be identified and weighed against potential benefits resulting from task variety, the use of different attention resources, and motivation. It is clear that ways of supporting operator detection need to be identified in order to improve detection rates. In addition to improving the selection and training of operators, ways of enhancing job and technology design need to be identified so that the human side of CCTV systems is better supported.

A CCTV surveillance typology would also assist in evaluating the potential effects of automated real time machine intelligence in CCTV systems. Automation is increasingly being built into surveillance systems, albeit to differing degrees (Donald, 2001). Digital systems have brought the capability for machine intelligence, but much of their potential has yet to be realised. Where machine intelligence is used to provide information on significant events, operators are still required to interpret the information and respond to it. A typology of CCTV surveillance tasks would provide a framework for understanding how machine intelligence alters vigilance demands.

Differences between participants with different work exposure indicate that future research on vigilance tasks related to jobs in the real world would do well to consider the sorts of samples used as this influences ecological validity. Where research on real life vigilance tasks is used, the sample should ideally consist of people performing these jobs. However, practical demands such as the number of people employed in these jobs and time off work create additional challenges in achieving this (Craig, 1984). Where different samples are used, caution should be exercised in making assumptions about ecological validity (Rosnow & Rosenthal, 1996).

8.6. Conclusion

The current research provides greater insight into vigilance processes in complex, real world tasks where information processing demands are high, and visual displays are dynamic and change on a continual basis. The original hypothesis that IGOs will help manage performance over time was not demonstrated in the research. However, it is clear that the area of vigilance and its dynamics in real world environments are more complex than expected by this researcher and the research literature in general. The current study has provided insights into a number of issues that related to an understanding of why this hypothesis was not confirmed, has highlighted important factors that should guide an understanding of vigilance research, and has provided some important insights into the directions where vigilance research should focus in the future.

The effectiveness of the IGO intervention in enhancing the detection of actual significant events was not demonstrated. However, qualitative analyses suggest that there were some indications (although statistically non-significant) that the intervention has some benefits. This illustrates the difficulty of inserting IGOs into an environment where visual stimuli are dynamic, change perspective and have different depths of field. However, there appears to be potential for IGOs to contribute to the detection of significant events with further research in the future. The identification of semantic distance (Isherwood et al., 2007) as an important characteristic of IGOs was confirmed through qualitative analyses. This is an important characteristic that has not previously been considered in the few studies on IGOs in CCTV (Andrew et al., 2003; Neil et al., 2007). It provides the basis from which to move forward with further research on IGOs. Future research into the other IGO characteristics that were identified and different IGO implementation strategies would contribute to a fuller understanding of the attention processes involved in visual searching.

In view of the complexity of vigilance processes, it has long been acknowledged that a single intervention is unlikely to address all vigilance issues (Craig, 1984). Therefore a systemic approach to developing a set of interventions that are inter-related and which reinforce each other is likely to prove beneficial in future. These include, for example, improved operator selection processes, training aimed at developing SA and broad frames of reference, visual analysis skills, control room and job design, and performance management systems. Based on Endsley's (2000) levels of SA, such training could be aimed at improving the perception and comprehension of events, as well as the ability to predict significant events.

The research has caused a fundamental reappraisal of the nature of vigilance under realistic work task conditions, and whether different work settings can be compared with each other (e.g., CCTV and X-ray baggage screening). The timing and variability of individual responses and changes in performance levels indicate that it is difficult to predict vigilance decrements over short periods of time. Similarly, the Andrew et al. (2003) study found that a vigilance decrement only occurred after two hours. Therefore vigilance decrements over far more extended time periods clearly need to be examined. Further studies of vigilance are going to have to focus on more realistic tasks where broad frames of reference and attention sets are integral parts of job performance and use samples that are appropriate to real world settings.

Previous vigilance research has tended to focus on the vigilance decrement rather than the overall level of vigilance performance. The current research clearly demonstrates that the overall level of vigilance is as much, if not more, of an issue than the traditional vigilance decrement. Low detection rates were evident throughout the designated ninety-minute period and only 50% (global sample) to 55% (operators only) of TBs were detected, indicating that a large proportion of TBs were missed. No participants detected all the TBs. This occurred despite the fact that all TBs were visible and no inferences were required for detection. This shows the difficulty of CCTV surveillance and detection, and the need to re-evaluate expectations of operator performance in the CCTV context. Besides CCTV surveillance, this may have implications for other work functions involving ongoing visual analysis and constant observation. At a theoretical level, it highlights the importance of the overall level of vigilance in complex and dynamic vigilance tasks. A low baseline detection rate contributes significantly to poor performance, regardless of whether a vigilance decrement occurs or not. The study also adds to an understanding of the processes involved in the overall level of vigilance, such as fluctuations in attention and task engagement levels. Previous researchers have not considered vigilance performance in this light.

Previous research has focused on identifying factors that modulate a steady and progressive linear type of vigilance decrement and have largely ignored individual variations in detection patterns. The current research contributes by identifying highly individualised responses to the demands of a complex and dynamic vigilance task. Decreases in detection levels begin at different times for different individuals, and some participants were better able to sustain attention than others. The response patterns were related to moment-by-moment fluctuations in task disengagement. These fluctuations have not previously been examined in a complex vigilance task and show that vigilance performance is frequently unstable and subject to continuous fluctuations over time, involving series of decreases and increases in detection rates. Performance fluctuations were observed both in individuals who experienced a traditional vigilance decrement and in those who did not show a decrement. The alternating series of performance fluctuations emphasise the need to move away from the assumption that the steady, linear vigilance decrement is the key performance issue in complex vigilance tasks.

The identification of work exposure as a variable that influences detection rates and the existence of both vigilance decrements and increments in the designated task are important contributions to knowledge. The absence of a vigilance decrement and existence of an increment in specialists, compared with novices and generalists, contrasts with most previous vigilance research. The differences were explained through specialists' broader frames of reference, higher levels of SA, their active approach to the searching process, and a greater ability to intentionally regulate their levels of task engagement. This is an important theoretical contribution because it provides a better understanding of the detection and searching processes.

The current research contributes to an understanding of the attention and searching processes in a complex and dynamic task. These processes form the centre of the CCTV surveillance process, yet knowledge of how they occur has been virtually non-existent in the past. This research has identified and explained a number of processes at the heart of visual searching and detection, including the role of broad frames of reference and attention sets, visual analysis, active search and task engagement. These assist in explaining the role of goal-directed attention guidance in search processes and have not previously been examined in the context of CCTV surveillance. The identification of the role of active searching is unique to this research. There are indications that SA is an important part of the detection process with CCTV and assists in explaining the perception, comprehension and prediction of significant events.

The participation of operators in this study represents a significant departure from the samples used in most vigilance research that have consisted of convenience samples of students. The use of a different type of task and sample have led to some different results being obtained from traditional vigilance studies and therefore questions the applicability of much previous vigilance research to complex and dynamic real world tasks. In addition, the identification of the importance of participants' work exposure is a variable that has not previously been examined for vigilance tasks of this nature.

The CCTV environment proved to be a rich context that allowed for the examination of vigilance, attention, and visual searching process in a dynamic environment. As expected, the vigilance dynamics and detection process are different in this environment from laboratory experiments using simple stimuli. The nature of the displays is important in contributing to vigilance dynamics and detection rates. This includes the continuous video stream rather than discrete events, the ongoing changes to the scenes depicted, the dynamic nature of the significant events, the dense clutter of the displays and meaning imbued in the scenes. These factors significantly alter the nature of the detection task and the vigilance dynamics. More specifically, the focus on operators assists in addressing the human side of CCTV systems which has been neglected in most previous research (Donald, C., 2008; Keval & Sasse 2006; Neil, et al., 2007; Wells et al., 2006). This has implications for much of the new imaging technology used in security, mapping and defence, where high levels of visual analysis and interpretation are required for them to be used effectively. This includes imaging technologies such as X-ray (goods and full body), infrared, thermal, terahertz, passive millimetre wave radar and active radar (Binstock & Minukas, 2010). The research has contributed to a deeper understanding of the nature of visual analysis in image intensive environments and how this might vary across different technologies.

References

ABI Research. (2008). *Video Surveillance: A market poised for $46 billion of explosive growth*. Retrieved from http://www.abiresearch.com/abiprdisplay.jsp?pressid=1081.

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. *Psychological Science*, *14*(5), 427-432.

Adams, J. A. (1987). Criticisms of vigilance research: A discussion. *Human Factors*, *29*(6), 737-740.

Adams, J. A., Humes, J. M., & Stenson, H. H. (1962). Monitoring of complex visual displays III: Effects of repeated sessions on human vigilance. *Human Factors*, *4*, 149-158.

Adams, J. A., Stenson, H. H., & Humes, J. M. (1961). Monitoring of complex visual displays II: Effects of visual load and response complexity on human vigilance. *Human Factors, 3,* 213-221.

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1991). *Strategic workload and the cognitive management of advanced multi-task system*. Report No. ADA259357, Cambridge MA: BBN Labs Inc.

Akerboom, S. P. (1994). 'Not suitable for children under 3 years of age': The effects of pictogram characteristics and consumer's attributes on comprehension of and intended compliance with the warning. *International Journal of Injury Control and Safety Promotion*, *1(*3), 175-188.

Aldridge, J. (1994). CCTV operational requirements manual, version 3. *Publication No. 17/94, Home Office, Police Scientific Development Branch.* Woodcock Hill, Sandridge, St Albans, Herts,UK.

Alluisi, E. A., Coates, G. D., & Morgan, B. B. (1976). Effects of temporal stresses on vigilance and information processing. Paper presented at NATO International Symposium on Vigilance, St Vincent, Italy, 3-6 August.

Andrew, C., Landgrebe, T., & Donald, C. (2003). *Enhanced surveillance platform validation testing: Methodology report.* Technical report, de Beers, South Africa.

Arnott, S.R., & Pratt, J. (2002). A review of attentional capture: On its automaticity and sensitivity to endogenous control: Commentary: Filter or disengagement? *Psicologica, 23*(2), 311-313.

Altman, D. G. (1991). *Practical Statistics for Medical Research*. London: Chapman and Hall.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception and Psychophysics*, 55, 485-496.

Baker, C. H. (1960). Maintaining the level of vigilance by means of artificial signals. *Journal of Applied Psychology*, *44*, 336-338.

Baker, C. H. (1961). Maintaining the level of vigilance by means of knowledge of results about a secondary vigilance task. *Ergonomics*, *4*, 311-316.

Baker, (1962). *Man and radar displays.* New York: MacMillan.

Balakrishnan, J. D. (1998). Measures and interpretations of vigilance performance - evidence against the detection criterion. *Human Factors*, *40*(4), 601-623.

Ballard, J. C. (1996). Computerized assessment of sustained attention: A review of factors affecting vigilance performance. *Journal of Clinical and Experimental Neuropsychology*, *18*(6), 843-863.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65-94.

Bauer, B., Jolicoeur, P., & Cowan, W. B. (1996a). Visual search for colour targets that are or are not linearly separable from distracters. *Vision Research*, *36*(10), 1439-1466.

Bauer, B., Jolicoeur, P., & Cowan, W. B. (1996b). Distractor heterogeneity versus linear separability in colour visual search. *Perception, 25*(11), 1281-1293.

Becker, A. B., Warm, J. S., Dember, W. N., & Hancock, P. A. (1991). Effects of feedback on perceived workload in vigilance performance. *Proceedings of the Human Factors Society 35th Annual Meeting (pp. 1491-1494), Santa Monica, CA: Human Factors and Ergonomics Society.*

Benner, P. (1984). *From novice to expert: Excellence and power in clinical nursing practice*. Menlo Park, CA: Addison-Wesley.

Bennett, T., & Fobes, J. L. (1994). *Screener proficiency and evaluation reporting system (SPEARS): human factors functional requirements and procedures.* FAA Technical Center. Report number: DOT/FAA/CT-94/65.

Bergum, B. O., & Lehr, D. J. (1963). End spurt in vigilance. *Journal of Experimental Psychology, 66*, 383-385.

Binstock, J., & Minukas, M. (2010). *Developing an operational and tactical methodology for incorporating existing technologies to produce the highest probability of an individual wearing an IED.* Unpublished Master's thesis, Naval Postgraduate School, Monterey, California, US.

Blankenberger, S., & Hahn, K. (1991). Effects of icon design on human-computer interaction. *International Journal of Man-Machine Studies*, *35*(3), 363-377.

Boles, D. B. (2002). Lateralized spatial processes and their lexical implications. N*europsychologia, 40,* 2125-2135.

Boles, D. B., & Law, M. B. (1998). A simultaneous task comparison of differentiated and undifferentiated hemispheric resource theories. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 204–215.

Boot, W. R., Brockmole J. R., & Simons, D. J. (2005). Attention capture is modulated in dual-task situations. *Psychonomic Bulletin & Review*, *12*(4), 662-668.

Borell-Carrió, F., & Epstein, R. M. (2004). Preventing errors in clinical practice: A call for self-awareness. *Annals of Family Medicine, 2*, 310-316.

Bravo, M. J., & Farid, H. (2006). Object recognition in dense clutter. *Perception and Psychophysics*, *68*(6), 911-918.

Bravo, M., & Nakayama, K. (1992). The role of attention in different visual search tasks. *Perception and Psychophysics*, *51*, 465-472.

Broadbent, D. E. (1953). Noise, paced performance and vigilance tasks. *British Journal of Psychology*, *44*, 295-303.

Broadbent, D.E. (1958). *Perception and communication.* London: Pergamon Press.

Broadbent, D. E., & Gregory, M. (1963). Division of attention and the decision theory of signal detection. *Proceedings of the Royal Society*, *158B*, 221-231.

Brockmole, J. R., & Henderson, J. M. (2005a). Object appearance, disappearance, and attention prioritization in real-world scenes. *Psychonomic Bulletin & Review*, *12*(6), 1061-1067.

Brockmole, J. R., & Henderson, J. M. (2005b). Prioritization of new objects in real-world scenes: Evidence from eye movements. *Journal of Experimental Psychology*, *31*(5), 857-868.

Buck, L. (1966). Reaction time as a measure of perceptual vigilance. *Psychological Bulletin, 65*(5), 291-304.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*, 523-547.

Butter, C. M. (1987). Varieties of attention and disturbances of attention: A neuropsychological analysis. In M. Jeannerod (Ed.), *Neurophysiological and neuropsychological aspects of spatial neglect* (pp. 1-24). Amsterdam: Elsevier Science Publishers.

Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement. *Psychonomic Bulletin & Review, 11*(5), 932-937.

Cameron, A., Kolodinski, E., May, H., & Williams, N. (2008). *Measuring the effects of video surveillance on crime in Los Angeles*. Technical report number CRB-08-00, prepared for the California Research Bureau by the USC School of Policy, Planning and Development, Los Angeles. Retrieved from http://www.library.ca.gov/crb/08/08-007.pdf

Carli, V. (2008). *Assessing CCTV as an effective safety and management tool for crime-solving, prevention and reduction*. International centre for the prevention of crime. Retrieved from http://www.crime-prevention-intl.org/uploads/media/pub_213_1.pdf

Carriére, J. S. A., Cheyne, J. A., & Smilek, D. (2008). Everyday attention lapses and memory failures: The affective consequences of mindlessness. *Consciousness and Cognition, 17*(3), 835-847.

Carroll, J. M., & Olson, J. (1987). *Mental models in human-computer interaction: Research issues about what the user of software knows*. Washington, DC: National Academy Press.

Castelhano, M. S., & Heaven, C. (2010). The relative contribution of scene context and target features to visual search in scenes. *Attention, Perception, & Psychophysics, 72*, 1283–1297.

Castelhano, M. S., & Henderson, J. M. (2008). Stable individual differences across images in human saccadic eye movements. *Canadian Journal of Experimental Psychology, 62*(1), 1-14.

Cavanagh, P., Arguin, M., & Treisman, A. (1990). Effect of surface medium on visual search for orientation and size features. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 479-491.

Chastain, G., Cheal, M., & Kuskova, V. (2002). Inappropriate capture by diversionary dynamic elements. *Visual Cognition, 9,* 355-381.

Chen, X., & Zelinsky, G. J. (2006). Real-world visual search is dominated by top-down guidance. *Vision Research*, *46*(24), 4118-4133.

Cheyne, J. A., Carriere, J. S. A., & Smilek, D. (2006). Absent-mindedness: Lapses of conscious awareness and everyday cognitive failures. *Consciousness and Cognition, 15*(3), 578-592.

Cheyne, J. A., Solman, G. J. F., Carriére, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition, 1*(11), 98-113.

Childs, J. M. (1976). Signal complexity, response complexity, and signal specification in vigilance. *Human Factors*, *18*, 149-160.

Cho, H., Ishida, T., Oyama, S., Inaba, R., & Takasaki, T. (2008). Assisting pictogram selection with categorized semantic. *IEICE Transactions on Information and Systems (Part D), 91*(11), 2638-2638.

Clarke, G. M., & Cooke, D. (2004). A basic course in statistics (5th ed.). Chichester: John Wiley & Sons.

Clarke, R.V. (1997). *Situational crime prevention: Successful case studies* (2nd ed.). New York: Harrow & Heston.

Cole, G. G., & Liversedge, S. P. (2006). Change blindness and the primacy of object appearance. *Psychonomic Bulletin & Review*, *13*(4), 588-594.

Colquhoun, W. P., & Edwards, R. S. (1970). Practice effects on a visual vigilance task with and without search. *Human Factors, 12*(6), 537-545.

Conway, A. R. A., & Kane, M. J. (2001). Capacity, control & conflict: An individual differences perspective on attentional capture. In Folk, C. L., & Gibson, B. S. (Eds*). Attraction, distraction and action: Multiple perspectives on attentional capture* (pp. 349-372). Amsterdam: Elsevier.

Corkill, J. D. (2006). *Intelligence led CCTV surveillance operations empowering the operator value adding the operation*. Paper presented at Mass Transport Security Conference, 4 - 5 December, Sofitel Hotel, Brisbane.

Craig, A. (1977). Broadbent and Gregory revisited: Vigilance and statistical decision. *Human Factors*, *19*, 25-36.

Craig, A. (1984). Human engineering: The control of vigilance. In J. S. Warm (Ed.), *Sustained attention in human performance* (pp. 247-291). Chichester, England: Wiley.

Craig, A. (1987). Signal detection theory and probability matching apply to vigilance*. Human Factors*, *29*(6), 645-652.

Cunningham, S., Scerbo, W., & Freeman, F. (2000). The electrocortical correlates of daydreaming during vigilance tasks. *The Journal of Mental Imagery, 24,* 61–72.

Cutler, V., & Paddock, S. (2009). Use of threat image projection (TIP) to enhance security performance. *Proceedings of IEEE 43rd Annual 2009 International Carnahan Conference on Security Technology*, 5-8 Oct., Zurich, 46 – 51. Retrieved from http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5335565. DOI:10.1109/CCST.2009.5335565

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and scene perception. *Psychological Science, 15*, 559-564.

Davies, D.R., & Parasuraman, R. (1982). *The psychology of vigilance.* New York: Academic Press.

Deery, H. A. (1999). Hazard and risk perception among young novice drivers. *Journal of Safety Research, 30*(4), 225-236.

De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effect of scene context on object identification. *Psychological Research, 52*, 317-329.

Dekker, S., & Hollnagel, E. (2004). Human factors and folk models. *Cognition, Technology, and Work, 6*, 70-86.

Dember, W. N., & Warm, J. S. (1979). *Psychology of perception* (2nd ed.). New York: Holt, Rinehart and Winston.

Department for Transport (2003). The selection of X-ray screeners. Transport Security Division, Technical Information Notice 1/03, United Kingdom.

Derryberry, D., & *Reed*, M. A. (*2001*). A multidisciplinary perspective on attentional control. In C. Folk & B. Gibson (Eds.), *Attraction, distraction, and action: Multiple perspectives on attentional capture* (pp. 325- 347). Amsterdam: Elsevier Science

Deutsch, G., Papanicolau, A. C., Bourbon, W. T., & Eisenberg, H. M. (1987). Cerebral Blood Flow Evidence of Right Frontal Activation in Attention Demanding Tasks. *International Journal of Neuroscience, 36*(1-2), 23-28.

Dickman*,* S.J. (2002). Dimensions of arousal*:* Wakefulness and vigor. *Human Factors, 44,* 429–442.

Di Stasi, L. L., Contreras, D., Cándido, A., Cañas, J. J., & Catena, A. (2011). Behavioral and eye-movement measures to track improvements in driving skills of vulnerable road users: First-time motorcycle riders. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*, 26-35.

Dockree, P. M., Kelly, S. P., Robertson, I. H., Reilly, R. B., & Foxe, J. J. (2005). Neurophysiological markers of alert responding during goal-directed behavior: a high-density electrical mapping study. *NeuroImage, 27*(3), 587-601.

Donald, C. (1998a). *Technical report on the validity of the surveillance and monitoring assessment exercise (SAMAE).* Technical Report No. 2/1998, Leaderware, SA. Retrieved from http://www.leaderware.com/SAMAE%20Technical%20Validation%20Study.htm

Donald, C. (1998b). Never underestimate the human factor. *International Security Review, 104*.

Donald, C. (2001). Vigilance. In J. Noyes, & M. Bransby (Eds), *People in control: Human factors in control room design.* London: The Institution of Electrical Engineers.

Donald, C. (2008). A strategic approach to security technology implementation - the case for CCTV and implications for public protection. *Key Note Address in Proceedings of the First Security Australia (SECAU) Conference*, 1 December, Edith Cowan University, Perth, Australia.

Donald, C., & Andrew, C. (2003). *Technology enhancement of CCTV operator effectiveness*. Paper presented at the CCTV User Group conference, Bolton, 9 April, 2003.

Donald, C., Andrew, C., & Landgrebe, T. (2007). *The Impact of CCTV TIP Implementation*. Paper presented at the Human Factors in Transport Security Academic Forum, Trinity College, Oxford, 24-25 September.

Donald, F. (2008). The classification of vigilance tasks in the real world. *Ergonomics*, *51*(11), 1643-1655.

Donald, F. (2010). A model of CCTV surveillance operator performance, *Ergonomics SA, 22*(2), 2-13.

Donald, F., & Donald, C. (2008). Vigilance and the implications of using threat image projection (TIP) for CCTV surveillance operators. *Proceedings of the First Security Australia (SECAU) Congress*, Security Research Centre, 1 December, Edith Cowan University, Perth, Australia.

Dowse, R., & Ehlers, M. S. (2004). Pictograms for conveying medicine instructions: Comprehension in various South African language groups. *South African Journal of Science*, *100*(11), 687-693.

Dreyfus, S. E., & Dreyfus, H. L. (1980). A five-stage model of the mental activities involved in directed skill acquisition. Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc.pdf

Dubbeld, L. (2005). The role of technology in shaping CCTV surveillance practices. *Information, Communication and Society, 8*(1), 84-100.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*, 501-517.

D’zmura, M. (1991). Color in visual search. *Vision Research, 31*, 951-966.

Edin, F., Klingberg, T., Johansson, P., McNab, F., Tegnér, J., & Compte, A. (2009). Mechanism for top-down control of working memory. *Proceedings of the National Academy of Science of the United States of America, 106*(16), 6802-6807. DOI: 10.1073/pnas.0901894106

Edkins, G. D., & Pollock, C. M. (1997). The influence of sustained attention on railway accidents. *Accident Analysis and Prevention*, *29*(4), 533-539.

Egeth, H. E., & Yantis, S. (1997). Visual attention; control, representation, and time course. *Annual Review of Psychology*, *48*(1), 269-298.

Ehinger, K., & Brockmole, J. R. (2008). The role of color in visual search in real-world scenes: Evidence from contextual cuing. *Perception and Psychophysics*, *70*(7), 1366-1378.

Eimer, M., Nattkemper, D., Schröger, E., & Prinz, W. (1996). Involuntary attention. In O. Neumann, & A. F. Sanders, (Eds.). *Handbook of perception and action, Vol. 3: Attention*, (pp. 155-184). London: Academic Press.

Einhaüser, W., Rutishauser, U., & Koch, C. (2008). Task demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision*, *8*(2), 1-19.

Endsley, M.R. (1995). Towards a theory of situation awareness in dynamic systems. *Human Factors*, *37*(1), 32-64.

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M. R. (2001). Designing for situation awareness in complex systems. *Proceedings of the Second International Workshop on Symbiosis of Humans, Artifacts, and Environment*, 26-27 November, Kyoto, Japan.

English, I. (2008). Intuition as a function of the expert nurse: A critique of Benner's novice to expert model. *Journal of Advanced Nursing, 18*(2), 387-393.

Ericsson, K. A., Charness, N., Hoffman, R. R., & Feltovich, P. J. (2006). *The Cambridge handbook of expertise and expert performance.* Cambridge: Cambridge University Press.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as raw data.* Cambridge, MA: MIT Press.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1476-1492.

Familant, M., & Detweiler, M. (1993). Iconic reference: evolving perspectives and an organising framework. *International Journal of Man-Machine Studies, 39*, 705-728.

Federico, P. A. (1995). Expert and novice recognition of similar situations. *Human Factors*, *37*(1), 105-122.

Findlay, J. M. (1997). Saccade target selection during visual search. *Vision Research*, *37*(5), 617-631.

Flach, J. M. (1995). Situation awareness: Proceed with caution. *Human Factors, 37*(1), 149-157.

Fleetwood, M. D., & Byrne, M. D. (2006). Modeling the visual search of displays: A revised ACT-R model of icon search based on eye-tracking data. In F. Detje, D. Dörner, & H. Schaub (Eds.), *Proceedings of the Fifth International Conference on Cognitive Modeling* (pp. 87-92). Bamberg, Germany: Universitas-Verlag Bamberg.

Folk, C. L., & Remington, R. W. (1996). When knowledge does not help: Limitation on the flexibility of attentional control. In A.F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp 77-106). Washington, DC: American Psychological Association.

Folk, C. L., Remington, R., & Johnston, J. C. (1992). Involuntary covert orienting is contingent upon attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1030-1044.

Folk, C. L., Remington, R., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 317-329.

Forster, S., & Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition, 111*, 345–355.

Forsythe, A., Mulhern, G., & Sawey, M. (2008). Confounds in pictorial sets: The role of complexity and familiarity in basic-level picture processing. *Behavior Research Methods*, *40*(1), 116-129.

Forsythe, A., Sheehy, N., & Sawey, M. (2003). Measuring icon complexity: An automated analysis. *Behavior Research Methods, Instruments, & Computers*, *35*(2), 334-342.

Foulkes, J., Crombie, N., McClumpha, A. J., & Shadrake, R. (1997). *The utility of X-ray image enhancement functions used by UK aviation security screeners.* Technical report number DERA/CHS/HS3/CR97146/1.0, DERA UK.

Foulsham, T. & Underwood, G. (2007). Can the purpose of inspection influence the potency of visual saliency in scene perception? *Perception*, *36*, 1123-1138.

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision, 8*(2), 1-17.

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception and Psychophysics*, *65*, 999-1010.

Frost & Sullivan (2008*). World video surveillance market – investment analysis*. Frost & Sullivan Research Service. Retrieved from http://www.frost.com/prod/servlet/report-brochure.pag?id=N396-01-00-00-00

Gale, A., & Christie, B. (1987). *Psychophysiology and the electronic workplace*. Chichester: John Wiley & Sons Ltd.

Galinsky, T. L., Rosa, R. R., Warm, J. S., & Dember, W. N. (1993). Psychophysical determinants of stress in sustained attention, *Human Factors, 35*(4), 603-614.

García, M., Badre, A. N., & Stasko, J. T. (1994). Development and validation of icons varying in their abstractness. *Interacting with Computers, 6*(2), 191-211.

Garvey, W. D., Taylor, F. V., & Newlin, E. P. (1959). *The use of "artificial signals" to enhance monitoring performance.* USN Research Laboratory, Report No. 5269.

Gerrard, G., Parkins, G., Cunningham, I., Jones, W., Hill, S., & Douglas, S. (2007). *National CCTV Strategy*. Home Office, UK. Retrieved from http://www.statewatch.org/news/2007/nov/uk-national-cctv-strategy.pdf

Gibson, B. S., & Jiang, Y. (1998). Surprise! An unexpected color singleton does not capture attention in visual search. *Psychological Science, 9*, 176-182.

Gibson, B. S., & Peterson, M. A. (2001). Inattentional blindness and attentional capture: Evidence for attention-based theories of visual salience. In Folk, C. L., & Gibson, B. S. (Eds). *Attraction, distraction and action: Multiple perspectives on attentional capture*. Amsterdam: Elsevier.

Gilchrist, I., & Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology*, *10*(19), 1209-1212.

Gilden, D. L., & Hancock, H. (2007). Response variability in attention deficit disorders. *Psychological Science, 18*, 796-802.

Gill, M., & Spriggs, A. (2005). *Assessing the impact of CCTV*. Home Office Research Study No. 292. U.K.: Home Office Research, Development and Statistics Directorate.

Gittins, D. (1986). Icon-based human-computer interaction. *International Journal of Man-Machine Studies, 24*(6), 519-543.

Goldstein, I. L., Johnston, W. A., & Howell, W. C. (1969). Complex vigilance: Relevant and irrelevant signals. *Journal of Applied Psychology, 53*(1), 45-48.

Gordon, R. D. (2004). Attentional allocation during the perception of scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 760-777.

Gorshkov, D. (2008). *Setting standards for CCTV & video analytics for transport systems*. Paper presented at the CCTV User Group Autumn Conference, Four Pillars Hotel, Cotswold Water Park, UK.

Greenberg, D. F., & Roush, J. B. (2009). The effectiveness of an electronic security management system in a privately owned apartment complex. *Evaluation Review*, *33*, 3-26.

Greene, C. (2008). Partners in crime. *CCTV Image*, *27*(44). Retrieved from http://www.cctvimage.com/downloads/CCTV_Image_27_June_2008.pdf

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., Szalma, J. L., & Parasuraman, R. (2003). The vigilance decrement reflects limitations in effortful attention, not mindlessness. *Human Factors*, *45*, 349-359.

Hackman, J. R., Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance, 16*(2), 250-279.

Hancock, H. H., Rogers, W. A., Schroeder, D., & Fisk, A. D. (2004). Safety symbol comprehension: Effects of symbol type, familiarity, and age. *Human Factors*, *46*(2), 183-195.

Hancock, P. A., & Hart, S. G. (2002). Defeating terrorism: What can human factors / ergonomics offer? *Ergonomics in Design*, *10*(1), 6-16.

Hauland, G. (2008). Measuring individual and team situation awareness during planning tasks in traiing of en route air traffic control. *International Journal of Aviation Psychology, 18*(3), 290-304.

He, J., Becic, E., Lee, Y., Jason, S., McCarley, J. (2011). Mind Wandering Behind the Wheel: Performance and Oculomotor Correlates. *Human Factors, 53*(1), 13-21.

Head, H. (1926). *Aphasia and kindred disorders of speech*. New York: Macmillan.

Helton, W. S., Dember, W. N., Warm, J. S., & Matthews, G. (1999). Optimism, pessimism, and false failure feedback: Effects on vigilance performance. *Current Psychology, 18,* 311-326.

Helton, W. S., Hollander, T. D., Warm, J. S., Matthews, G., Dember, W.N., Wallart, M., Beauchamp, G., Parasuraman, R., & Hancock, P.A. (2005). Signal regularity and the mindlessness model of vigilance. *British Journal of Psychology*, *96*(2), 249-261.

Helton, W. S., Hollander, T. D., Warm, J. S., Tripp, L. D., Parsons, K., Matthews, G., Dember, W. N., Parasuraman, R., & Hancock, P. A. (2007). The abbreviated vigilance task and cerebral hemodynamics. *Journal of Clinical and Experimental Neuropsychology, 29*(5), 545-552.

Helton, W. S., Kern, R. P., & Walker, D. R. (2009). Conscious thought and the sustained attention to response task. *Consciousness and Cognition, 18*, 600-607.

Helton, W. S., Weil, L., Middlemiss, A., & Sawers, A. (2010). Global interference and spatial uncertainty in the Sustained Attention to Response Task (SART). *Consciousness and Cognition, 19,* 77–85.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498-504.

Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science*, *16*(4), 219-222.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye-movements during visual search in real-world scenes. In R. van Gompel, M. Fischer, W. Murray & R. Hill (Eds.), *Eye movement research: Insights into mind and brain* (pp. 537-562). Oxford: Elsevier.

Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 210-228.

Hillstrom, A. P., & Yantis, S. (1994). Visual motion and attentional capture. *Perception and Psychophysics, 55*, 399-411.

Hitchcock, E. M., Dember, W. N., Warm, J. S., Moroney, B. W., & See, J. E. (1999). Effects of cueing and knowledge of results on workload and boredom in sustained attention. *Human Factors, 41,* 365-372.

Hitchcock, E. M., Warm, J. S., Matthews, G., Dember, W. N., Shear, P. K., Tripp, L. D., et al. (2003). Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. *Theoretical Issues in Ergonomics Science, 4*(1-2), 89-112.

Hodsoll, J. P., Humphreys, G. W., & Braithwaite, J. J. (2006). Dissociating the effects of similarity, salience, and top-down processes in search for linearly separable size targets. *Perception and Psychophysics*, *68*(4), 558-570.

Holland, J. G. (1958). Human vigilance – the rate of observing an instrument is controlled by the schedule of signal detections. *Science*, *128*, 61-67.

Hollenbeck, J. R., Ilgen, D. R., Tuttle, D. B., & Sego, D. J. (1995). Team performance on monitoring tasks: An examination of decision errors in contexts requiring sustained attention. *Journal of Applied Psychology*, *80*(6), 685-696.

Hollingsworth, D. E., McAuliffe, S. P., & Knowlton, B. J. (2011). Temporal Allocation of Visual Attention in Adult Attention Deficit Hyperactivity Disorder. *Journal of Cognitive Neuroscience, 13*(3), 298-305.

Houts, P. S., Doak, C. C., Doak, L. G., & Loscalzo, M. J. (2006). The role of pictures in improving health communication: A review of research on attention, comprehension, recall, and adherence. *Patient Education and Counseling, 61*(2), 173-190.

Howell, W. C., Johnston, W. A., & Goldstein, I. L. (1966). Complex monitoring and its relation to the classical problem of vigilance. *Organizational Behavior & Human Performance*, *1*, 129-150.

Huey, B. M., & Wickens, C. D. (1993). *Workload transition: Implications for individual and team performance*. Washington, DC: National Academy Press.

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct manipulation interfaces. *Human-Computer Interaction, 1,* 311-338.

Ibarra-Orozco, R., Gonzalez-Mendoza, M., Hernandez-Gress, N., Diederichs, F., & Kortelainen, J. (2008). Towards a ready-to-use drivers' vigilance monitoring system. *Proceedings of the International Conference on Computational Intelligence for Modelling Control & Automation*, 10 – 12 December, Vienna, 802-807. DOI: 10.1109/CIMCA.2008.161.

Irwin, D. E., Colcombe, A. M., Kramer, A. F. & Hahn, S. (2000). Attentional and oculomotor capture by onset luminance and color singletons. *Vision Research*, *40*(10-12), 1443-1458.

Isherwood, S. J., McDougall, S. J. P., & Curry, M. B. (2007). Icon identification in context: The changing role of icon characteristics with user experience. *Human Factors*, *49*(3), 465.

Itti, L. (2005). Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition*, *12*(6), 1093-1123.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10-12), 1489-1506.

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How novices and experts interpert dynamic stimuli. *Learning and Instruction, 20*(2), 146-154.

Jerison, H. J. (1959). Effects of noise on human performance. *Journal of Applied Psychology, 43*(2), 96-101.

Jerison, H. J. (1977). Vigilance: Biology, psychology, theory and practice. In R. R. Mackie (Ed.), *Vigilance: Theory, operational performance, and physiological correlates* (pp. 27-40). New York: Plenum.

Johnson, A., & Proctor, R. W. (2004). *Attention: Theory and practice*. Thousand Oaks, CA: Sage.

Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin*, *120*(1), 113-139.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception and Psychophysics*, *43*, 346-354.

Kaar, M. (2007). *A critical investigation of the merits and drawbacks of in-depth interviews.* Norderstedt: GRIN Verlag.

Kahneman, D. (1973). *Attention and effort.* New York: Prentice Hall.

Kanfer, R., & Ackerman, P. L. (1996). A self-regulatroy skills perspective to reducing cognitive interference. In I. G. Sarason, & G. R. Pierce (Eds.), *Cognitive interference: Theories, methods, and findings* (pp. 153-172). Mahwah NJ: Lawrence Erlbaum.

Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. D. (2001). Comparison of expert and novice scan behaviors during VFR flight., 1-6. Retrieved from http://www.faa.gov/library/online_libraries/aerospace_medicine/sd/media/Kasarski.pdf

Keval, H., & Sasse, M. A. (2006). Man or gorilla? Performance issues with CCTV technology in security control rooms. *Proceedings of the 16th World Congress on Ergonomics Conference*, *International Ergonomics Association*, 10-14 July, Maastricht, Netherlands.

Keval, H., & Sasse, M. A. (2008). To catch a thief - you need at least 8 frames per second: The impact of frame rates on user performance in a CCTV detection task. *Proceedings of the International Multi-media Conference*, 26 – 31 October, Vancouver, British Columbia, Canada, 241-244.

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, *8*, 255-273.

Kirlik, A., & Strauss, R. (2006). Situation awareness as judgment 1: Statistical modeling and quantitative measurement. *International Journal of Industrial Ergonomics*, *36*(5), 463-474.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4*(4), 219-227.

Koelega, H. S. (1996). Sustained attention. In O. Neumann, & A. F. Sanders, (Eds.). *Handbook of perception and action, Vol. 3: Attention* (pp. 277- 331). London: Academic Press.

Kramer, A. F., Trejo, L. J., & Humphrey, D. G. (1996). Psychophysiological measures of workload: Potential applications to adaptively automated systems. In R. Parasuraman, & M. Mouloua (Eds), *Automation and human performance: Theory and applications* (pp. 137-181). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Lal, S. K. L., & Craig, A. (2002). Driver fatigue: Electroencephalography and psychological assessment. *Psychophysiology, 39*(2), 313-321.

Lanzetta, T. M., Dember, W. N., Warm, J. S., & Berch, D. B. (1987). Effects of task type and stimulus heterogeneity on the event rate function in sustained attention. *Human Factors, 29*(6), 625-633.

Laughery, K. R. (2006). Safety communications: Warnings*. Applied Ergonomics*, 37(4), 467-478.

Law, B., Atkins, M. S., Kirkpatrick, A. E., & Lomax, A. J. (2004). Eye gaze patterns differentiate novices and experts in a virtual laparoscopic surgery training environment. *Symposium on Eye Tracking Research Applications.* San Antonio, Texas, 41-48.

Leber, A. E., & Egeth, H. E. (2006). Attention on autopilot: Past experience and attentional set. *Visual Cognition*, *14*(4/5/6/7/8), 565-583.

Leblanc, E., & Jolicoeur, P. (2005). The time course of the contingent spatial blink. *Canadian Journal of Experimental Psychology*, *59*(2), 124-131.

Levy,F. (1980). The development of sustained attention (vigilance) and inhibition in children: Some normative data. *Journal of Child Psychology and Psychiatry, 21*(1), 77-84.

Li, F., van Rullen, R., Koch, C., & Perona, P. (2005). Why does natural scene categorization require little attention? Exploring attentional requirements for natural and synthetic stimuli. *Visual Cognition*, *12*(6), 893-924.

Litchfield, D., Ball, L. J., Donovan, T., Manning, D. J., & Crawford, T. (2010). Viewing another person's eye movements improves identification of pulmonary nodules in chest X-ray inspection. *Journal of Experimental Psychology: Applied, 16*(3), 251-262.

Liu, Y. (1996). Quantitative assessment of effects of visual scanning on concurrent task performance. *Ergonomics*, *39*(3), 382-399.

Lloyd-Jones, T. J., & Luckhurst, L. (2002). Effects of plane rotation, task and visual complexity on recognition of familiar and chimeric objects. *Memory and cognition*, *30*, 499-510.

Loeb, M., & Alluisi, E. A. (1984). Theories of vigilance. In J. S. Warm (Ed.), *Sustained attention in human performance* (pp. 247-291). Chichester, England: Wiley.

Long, C. M., & Waag, W. L. (1981). Limitations and practical applicability of d′ and ß as measures, *Human Factors*, *23*, 283-290.

Lorenz, B., Biella, M., Teegen, U., Stelling, D., Wenzel, J., Jakobi, J., et al. (2006). Performance, situation awareness, and visual scanning of pilots receiving onboard taxi navigation support during simulated airport surface operation. *Human Factors and Aerospace Safety, 6*(2), 135-154.

Lowe, V. J., Hofman, J. M., de Long, D. M., Patz, E. F., & Coleman, R. E. (1994). Semiquantitative and visual analysis of FDG-PET images in pulmonary abnormalities. *Journal of Nuclear Medicine, 35(*11), 1771-1776.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT Press.

Mackie, R. R. (1987). Vigilance research - are we ready for countermeasures? *Human Factors*, *29*(6), 707-723.

Mackie, R. R., Wylie, C. D., & Smith, M. J. (1994). Countering loss of vigilance in sonar watchstanding using signal injection and performance feedback, *Ergonomics*, *37*(7), 1157-1184. Retrieved from www.tandf.co.uk/journals/tf/00140139.html

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology, 1*, 6-21.

Mackworth, N. H. (1950). *Researches in the measurement of human performance*. Medical Research Council Special Report No. 268, HMSO, London.

Mackworth, N.H. (1957). Some factors affecting vigilance. *Symposium on Vigilance, /Advancement of Science*, 53, 389-393.

Mackworth, J. F. (1970). *Vigilance and attention*. Baltimore: Penguin.

MacLean, K., Aichele, S., Bridwell, D., Mangun, G., Wojciulik, E., & Saron, C. (2009). Interactions between endogenous and exogenous attention during vigilance. *Attention, Perception and Psychophysics, 71*(5), 1042-1058.

MacMillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge: Cambridge University Press.

Mansoor, L., & Dowse, R. (2007). Written medicines information for South African HIV/AIDS patients: Does it enhance understanding of co-trimoxazole therapy? *Health Education Research*, *22*(1), 37-48.

Martin-Emerson, R., & Kramer, A. F. (1997). Offset transients modulate attentional capture by sudden onsets. *Perception and Psychophysics*, *59*, 739-751.

Masalonis, A. J., & Parasuraman, R. (2003). Fuzzy signal detection theory: analysis of human and machine performance in air traffic control, and analytic considerations. *Ergonomics, 46*(11), 1045-1074.

Matthews, G., & Davies, D. R. (1998). Arousal and vigilance: Still vital at fifty. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1, 772-777.

Matthews, G., & Davies, D. R. (2001). Individual differences in energetic arousal and sustained attention: A dual-task study. *Personality and Individual Differences, 31*(4), 575-589.

Matthews, G., Davies, D. R., & Holley, P. (1993). Cognitive predictors of vigilance. *Human Factors, 35*(1), 3-24.

Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R. B. (2000). *Human performance, cognition, stress and individual differences*. East Sussex: Psychology Press.

Mazer, J. A., & Gallant, J. A. (2003). Goal-related activity in V4 during free viewing visual search. *Neuron, 40*(6), 1241-1250.

McCarley, J. S., Kramer, A. F., Wickens, C. D., Vidoni, E. D., & Boot, W. R. (2004). Visual skills in airport security inspection. *Psychological Science*, *15*, 302-306.

McDougall, S. J. P., Curry, M. B., & de Bruijn, O. (1999). Measuring symbol and icon characteristics: Norms for concreteness, complexity, meaningfulness. *Behavior Research Methods, Instruments, & Computers*, *31*(3), 487-519.

McDougall, S. J. P., Curry, M. B., & de Bruijn, O. (2001). The effects of visual information on users' mental models: An evaluation of pathfinder analysis as a measure of icon usability. *International Journal of Cognitive Ergonomics*, *5*(1), 59-84.

McDougall, S. J. P., & Isherwood, S. J. (2009). What's in a name? The role of graphics, functions, and their interrelationships in icon identification. *Behavior Research Methods*, *41*(2), 325-336.

McDougall, S., Tyrer, V., & Folkard, S. (2006). Searching for signs, symbols, and icons: Effects of time of day, visual complexity, and grouping. *Journal of Experimental Psychology: Applied*, *12*, 118-128.

McPeek, R. M., Maljkovic, V., & Nakayama, K. (1999). Saccades require focal attention and are facilitated by a short-term memory. *Vision Research*, *39*, 1555-1566.

McVay, J. C., Kane, M. J., Kwapil, T. R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological context. *Psychonomic Bulletin and Review, 16*(5), 857-863.

Messick, S. (1962). Response style and content measures from personality inventories. *Educational and Psychological Measurement, 22,* 41-56.

Methot, L. L., & Huitema, B. E. (1998). Effects of signal probability on individual differences in vigilance. *Human Factors*, *40*(1), 102-110.

Miller, L. A., & Stanney, K. M. (1997). The effect of pictogram-based interface design on human-computer performance. *International Journal of Human-Computer Interaction*, *9*(2), 119-132.

Milosevic, S. (1975). Changes in detection measures and skin resistance during an auditory vigilance task. *Ergonomics*, *18*, 1-18.

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors*, *38*(2), 311-322.

Moores, E., Laiti, L., & Chelazzi, L. (2003). Associative knowledge controls deployment of visual selective attention. *Nature Neuroscience*, *6*(2), 182-189.

Moray, N. (1984). Attention to dynamic visual displays in man-machine systems. In R. Parasuraman, & D. R. Davies (Eds.*), Varieties of attention* (pp. 485-514). Orlando: Academic Press.

Moray, N., & Haudegond, S. (1998). An absence of vigilance decrement in a complex dynamic task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Cognitive Engineering and Decision Making*, 234-238.

Mortier, K., Donk, M., & Theeuwes, J. (2003). Attentional capture within and between objects. *Acta Psychologica*, 113(2), 133-145.

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review*, *112*(1), 217-242.

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science*, *12*(1), 9-17.

Mouloua, M., & Parasuraman, R. (1995). Aging and cognitive vigilance: Effects of spatial uncertainty and event rate. *Experimental Aging Research*, *21*(1), 17-32.

Műller, C., & Boos, D. (2004). *Zurich main railway station: A typology of public CCTV systems*. Paper presented the Conference on CCTV and Social Control: The politics and practice of video surveillance – European and Global perspectives, 8 - 9 January, Sheffield, UK.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387-399.

Neiderman, E. C., & Fobes, J. L. (2005). *Threat image projection system*. United States Patent 6899540. Retrieved from www.freepatentsonline.com.

Neil, D., Thomas, N., & Baker, B. (2007). Threat image projection in CCTV. *Proceedings of the 41st Annual International IEEE Carnahan Conference on Security Technology*, Ottawa, ON, Canada, pp. 272-280.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology, 7*, 480-494.

Norris, C., & Armstrong, G. (1999). *The maximum surveillance society: The rise of CCTV*. Oxford: Berg.

Northern, J. J. (2010). *Anxiety and Cognitive Performance: A Test of Predictions Made by Cognitive Interference Theory and Attentional Control Theory*. Unpublished Masters thesis at Bowling Green State University, US.

Nuechterlein, K. H., Parasuraman, R., & Jiang, Q. (1983). Visual sustained attention: Image degradation produces rapid sensitivity decrement over time. *Science*, *220*(4594), 327-329.

Nunes, A., & Mogford, R. H. (2003). Identifying controller strategies that support the 'picture.' *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monica, CA: Human Factors and Ergonomics Society.

Oliva, A., Torralba, A., Castelhano, M. S., & Henderson, J. M. (2003). Top-down control of visual attention in object detection. *Proceedings of the International Conference on Image Processing*, 14-17 September, Barcelona, Spain, v. 1, pp. 253-256.

Olivers, C., Spalek, T., Kawahara, J., Di Lollo, V., Dux, P., Asplund, C., & Marois, R. (2009). The attentional blink: Increasing target salience provides no evidence for resource depletion. A commentary on Dux, Asplund, and Marois (2008): Both exogenous and endogenous target salience manipulations support resource depletion accounts of the attentional blink: A reply to Olivers, Spalek, Kawahara, and Di Lollo. *Psychonomic Bulletin & Review, 16*(1), 214-225.

Paivio, A. (1975). *Imagery and verbal processes*. New York: Holt, Reinhart & Winston.

Pappachan, P., & Ziefle, M. (2008). Cultural influences on the comprehensibility of icons in mobile-computer interaction. *Behaviour & Information Technology*, *27*(4), 331-337.

Parasuraman, R. (1976). Consistency of individual differences in human vigilance performance. *Journal of Applied Psychology*, *61*, 486-492.

Parasuraman, R. R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science*, *205*, 924-927.

Parasuraman, R. (1984). Sustained attention in detection and discrimination. In R. Parasuraman, & D. R. Davies (Eds.), *Varieties of attention* (pp. 243-272). Orlando: Academic Press.

Parasuraman, R. (1986). Vigilance, monitoring and search. In K. Boff, K. Kaufman and J. Thomas (Eds.), *Handbook of perception and human performance, Volume 2: Cognitive processes and performance.* Wiley: New York.

Parasuraman, R. (1998). The attentive brain: Issues and prospects. In R. Parasuraman (Ed.), *The attentive brain*, (pp. 3-15). Cambridge, MA: The MIT Press.

Parasuraman, R., & Davies, D. R. (1976). Decision theory analysis of response latencies in vigilance. *Journal of Experimental psychology: Human Perception and Performance*, 2, 569-582.

Parasuraman, R., & Davies, D. R. (1977). A taxonomic analysis of vigilance performance. In: R.R. Mackie (Ed.), *Vigilance: Theory, operational performance, and physiological correlates* (559–574). New York: Plenum.

Parasuraman, R., de Visser E., Clarke, E.,McGarry, R., Hussey, E., Shaw, T. & Thompson, J. C. (2009). Detecting threat-related intentional actions of others: Effects of image quality, response mode, and target cueing on vigilance. *Journal of Experimental Psychology: Applied*, *15*(4), 275-290.

Parasuraman, R., Mouloua, M. (1987). Interaction of signal discriminability and task type in vigilance decrement. *Perception & Psychophysics, 41*, 17–22.

Parasuraman, R., & Mouloua, M. (1996). *Automation and human performance: Theory and applications.* Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Parasuraman, R., Warm, J.S., & Dember, W. N. (1987). Vigilance: Taxonomy and utility. In L. S. Mark, J. S. Warm, & R. L. Huston (Eds.), *Ergonomics and human factors: Recent research*, (pp. 11-32). New York: Springer-Verlag.

Pardo, J. V., Fox, P. T., & Raichle, M. E. (1991). Localization of a human system for sustained attention by positron emission tomography. *Nature, 349*, 61-64.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual selective attention. *Vision Research*, *42*, 107-123.

Pashler, H. E. (1999). *The psychology of attention*. Cambridge, Massachussetts: The MIT Press.

Phillips, C. (1999). A review of CCTV evaluations: Crime reduction effects and attitudes towards its use. In K. Painter, & N. Tilley (Eds.), *Surveillance of public space: CCTV, street lighting and crime prevention* (pp. 123-156). Monsey: Criminal Justice Press.

Pigeau, R. A., Angus, R. G., O'Neill, P., & Mack, I. (1995). Vigilance latencies to aircraft detection among NORAD surveillance operators. *Human Factors*, *37*(3), 622-634.

Polson, M. C., & Friedman, A. (1988). Task-sharing within and between hemispheres. *Human Factors, 30*(5), 633-643.

Pomplun, M., Reingold, E.M., & Shen, J. (2003). Area activation: A computational model of saccadic selectivity in visual search. *Cognitive Science*, *27*(2), 299-312.

Pope, I. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology, 40*(1-2), 187-195.

Posner. M, I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25-42.

Proulx, M. J., & Egeth, H. E. (2006). Target-nontarget similarity modulates stimulus-driven control in visual search. *Psychonomic Bulletin and Review*, *13*(3), 524-530.

Rauschberger, R., & Yantis, S. (2006). Perceptual coding efficiency in visual search. *Journal of Experimental Psychology: General*, *135*(1), 116-131.

Remington, R. W., Folk, C. L., McLean, J. P. (2001). Contingent attentional capture or delayed allocation of attention? *Perception and Psychophysics*, *63*(2), 298-307.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, *53*, 245-277.

Richard, C. M., Wright, R. D., & Ward, L. M. (2003). Goal-driven modulation of stimulus-driven attentional capture in multiple-cue displays. *Perception and Psychophysics, 65*, 939-955.

Rijlaarsdam, G., van den Bergh, H., & Couzijn, M. (1996). *Theories, models and methodology in writing research*. Amsterdam: Amsterdam University Press.

Robertson, I. H. (2003). The absent mind attention and error. *The Psychologist, 16*(9), 476–479.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35*(6), 747-758.

Rogers, Y. (1989). Icons at the interface: Their usefulness. *Interacting with Computers*, *1*, 105-117.

Rolfs, M., Jonikaitas, D., Deubel, H., Cavanagh, P. (2010). Predictive remapping of attention across eye movements. *Nature Neuroscience, 14,* 252-256.

Rose, C. L., Murphy, L. B., Byard, L., & Nikzad, K. (2002). The role of the big five personality factors in vigilance performance and workload. *European Journal of Personality*, *16*, 185-207.

Rosenholtz, R. (1999). A simple saliency model predicts a number of motion popout phenomena. *Vision Research*, *39*, 3157-3163.

Rosenthal, R., & Rosnow, R. L. (2008). *Essentials of behavioural research: Methods and data analysis* (3rd ed.). Boston: McGraw Hill.

Rosnow, R. L., & Rosenthal, R. (1996). *Beginning behavioral research: A conceptual primer* (2nd ed.). Englewood Cliffs: Prentice Hall.

Rother, H. A. (2008). South African farm workers' interpretation of risk assessment data expressed as pictograms on pesticide labels. *Environmental Research, 108*(3), 419-427.

Sarason, I. G., Sarason, B.R., Keefe, D. W., Hayes, B.E., & Shearin, E. N. (1986). Cognitive interference: Situational determinants and trait like characteristics. *Journal of Personality and Social Psychology, 1,* 215–26.

SAS Enterprise Guide (n.d.). Retreived on 1 June 2011 from support.sas.com/documentation/onlinedoc/guide/

SAS Institute. (2008). *SAS Enterprise Guide version 4.2*. Cary, North Carolina, USA.

SAS Institute. (2008). SAS version 9.2. Cary, North Carolina, USA.

Sawin, D. A., & Scerbo, M. W. (1993). Vigilance: Where has all the workload gone? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Visual Performance,* 1383-1387.

Sawin, D. A., & Scerbo, M. W. (1994). Vigilance: How to do it and who should do it. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Visual Performance,* 1312-1316.

Sawin, D. A., & Scerbo, M. W. (1995). Effects of instruction type and boredom proneness in vigilance: implications for boredom and workload. *Human Factors*, *37*(4), 752-765.

Schmidke, I. I. (1976). Vigilance. In E. Simonson & P.C. Weiser (Eds.), *Psychological and physiological correlates of work and fatigue*. Springfield, MA: Thomas.

Schmidtke, H., & Micko, H. (1964). *Untersuchungen über die reaktionseit bei dauerbeobachtung.* Köln: Opladen

Security General. (2002). *Human factors in civil aviation security operations*. Technical Report No. 9808 AN/765, International Civil Aviation Organization, Canada.

Security Park. (2001). *Space-age surveillance comes to Newham*. Retrieved from www.securitypark.co.uk/pfv.asp?articleid=1832.

See, J.E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin*, *117*, 230-249.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in the human cortex. *Psychological Science*, *16*(2), 114-122.

Shiffrin, R. M. (1988). Attention. In R.C. Atkinson, R.J. Herrnstein, G. Lindzey, & R.D. Luce (Eds.), *Stevens' handbook of experimental psychology* (2nd ed., Vol. 2). New York: Wiley.

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences*, *4*(4), 147-155.

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

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*, 261-267.

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., et al. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition, 13*(4), 657-690.

Smallwood, J., Fishman, D. J., & Schooler, J. W. (2007). Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin and Review, 14*(2), 230-236.

Smallwood, J., Obonsawin, M., & Heim, D. (*2003*). Task unrelated thought: the role of distributed processing. *Consciousness and Cognition, 12,* 169-189.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*(6), 946-958.

Smit, A. S., Eling, P. A. T. M., & Coenen, A. M. L. (2004). Mental effort causes vigilance decrease due to resource depletion. *Acta Psychologica*, *115*, 35-42.

Smith, G. J. D. (2004). Behind the scenes: Examining constructions of deviance and informal practices among CCTV control room operators in the UK. *Surveillance and Society*, *2*(2/3), 376-395.

Smith, R. L., & Lucaccini, L. F. (1969). Vigilance research: Its application to industrial problems. *Human Factors, 11*, 149-156.

Snodgrass, J. G., & Vanderwart, M. (1980). Standardized set of 260 pictures: Norms of name agreement, usage agreement, word familiarity, and visual complexity. *Journal of Experimental and Human Learning and Memory*, *6*, 174-215.

Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(2), 248-261.

South African Concise Oxford Dictionary. (2002). Oxford: Oxford University Press.

Squires, P. (2010). *Evaluating CCTV: Lessons from a surveillance culture*. Paper presented at the Final Conference of the European Forum for Urban Safety: Citizens, Cities and Video-Surveillance Programme and for the EFUS CCTV Charter launched in Rotterdam, May 2010. Retrieved from http://www.fesu.org/

Suzuki, S., & Cavanagh, P. (1995). Facial organization blocks access to low level features: An object inferiority effect. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 901-913.

Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., Meier, A., & Eggemeier, F. T. (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors*, *46*(2), 219-233.

Temple, J. G., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., & Matthews, G. (2000). The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Human Factors*, *42*, 183-194.

Thatcher, A., Mahlangu, S., & Zimmerman, C. (2006). Accessibility of ATMs for the functionally illiterate through icon-based interfaces. *Behaviour & Information Technology*, *25*(1), 65-81.

Thayer, R. E. (1989). *The biopsychology of mood and arousal*. Oxford: Oxford University Press.

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception and Psychophysics*, *50*, 184-193.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception and Psychophysics*, *51*, 599-606.

Theeuwes, J. (1993). Visual selective attention: A theoretical analysis. *Acta Psychologica*, *83*, 93-154.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 799-806.

Theeuwes, J. (1996). Visual search at intersections: An eye-movement analysis. In A. G. Gale, I. D. Brown, S. P. Taylor, & C. M. Haslegrave (Eds.), *Vision in Vehicles 5* (pp. 203-212). Amsterdam: Elsevier.

Theeuwes, J., & Godijn, R.J. (2002). Irrelevant singletons capture attention: Evidence from inhibition of return. *Perception and Psychophysics*, *64*(5), 764-770.

Thiel, G. (2000). Automatic CCTV Surveillance - Towards the VIRTUAL GUARD. *IEEE AES Systems Magazine, July,* 3-9.

Tien, G., Zheng, B., Swindells, C., & Atkins, M. S. (2010). Measuring situation awareness of surgeons in laparoscopic training. *Proceedings of Eye Tracking Research and Applications,* 149-152.

Tomporowski, P. D., & Tinsley, V. F. (1996). Effects of memory demand and motivation on sustained attention in young and older adults. *The American Journal of Psychology*, *109*(2), 187-195.

Torralba, A., Castelhano, M. S., Henderson, J. M., & Oliva, A. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766-789.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *21*(1), 97-136.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence form search asymmetries. *Psychological Review*, *95*, 15-48.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 459-478.

Tullio, J., Huang, E., Wheatley, D., Zhang, H., Guerrero, C., & Tamdoo, A. (2010). Experience, adjustment, and engagement: The role of video in law enforcement. Proceedings of CHI, 10-15 April, Atlanta, Georgia, USA, 1505-1514.

Turatto, M., & Galfano, G. (2000). Color, form and luminance capture visual attention in visual search. *Vision Research*, *40*(13), 1639-1643.

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology, 59* (11), 1931 – 1949.

Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map. *European Journal of Cognitive Psychology*, *18*, 321-343.

Underwood, G., Foulsham, T., van Loon, E., & Underwood, J. (2005). Visual attention, visual saliency, and eye movements during the inspection of natural scenes. In J. Mira & J. R. Alvarez (Eds.) *Proceedings of Artificial intelligence and knowledge engineering applications: A bioinspired approach*, *pt 2, 3562*, 459-468.

van Voorthuijsen, G., van Hoof, H., Klima, M., Roubik, K., Bernas, M., & Pata, P. (2005). CCTV effectiveness study. *Proceedings of the 39th IEEE Annual 2005 International Carnahan Conference on Security Technology*, Piscataway, 105-108.

Van Wert, M., Horowitz, T. S., & Wolfe, J. M. (2009). Even in correctable search, some types of rare targets are frequently missed. *Attention, Perception and Psychophysics*, *71*(3), 541-553.

van Zoest, W., Giesbrecht, B., Enns, J. T., Kingstone, A. (2006). New reflections on visual search: Interitem symmetry matters! *Psychological Science*, *17*(6), 535–542.

Vecera, S., & Farah, M. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1-14.

Walker, G. H., Stanton, N. A., Salmon, P. M., Jenkins, D. P., Rafferty, L., & Ladva, D. (2010). Same or different? Generalising from novices and experts in military command and control studies. *International Journal of Industrial Ergonomics, 40*(5), 473-483.

Wallace, E., Diffley, C., Baines, E. & Aldridge, J. (n.d.) *Ergonomic Design Considerations for Public Area CCTV Safety and Security Applications*. Home Office, Police Scientific Development Branch (PSDB), UK.

Wallis, D., & Newton, G. de C. (1957). *An experiment on synthetic signal injection as a method of improving radar detection.* Technical Report No. W2/4, Department of Sensory Psychology, The Admiralty, UK.

Warm, J. S. (1977). Psychological processes in sustained attention. In R. R. Mackie (Ed.). *Vigilance: Theory, operational performance, and physiological correlates* (pp. 623-644). New York: Plenum.

Warm, J. S. (1984). *Sustained attention in human performance*. Wiley: Chichester.

Warm, J. S., & Dember, W. N. (1998). Tests of vigilance taxonomy. In Hoffman, R. R., Sherrick, M. F., & Warm, J.S. (Eds.). *Viewing psychology as a whole: The integrative science of William N. Dember* (pp. 87-112). Washington, DC: American Psychological Association.

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In R. Parasuraman, & M. Mouloua (Eds.). *Automation and human performance: Theory and applications. Human factors in transportation* (pp 183-200). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, *50*(3), 433-441.

Wells, H., Allard, T., & Wilson, P. (2006). *Crime and CCTV in Australia: Understanding the relationship*. Centre of Applied Psychology and Criminology, Faculty of Humanities and Social Sciences, Bond University, Gold Coast, Australia.

Welsh, B. C., & Farrington, D. P. (2002). *Crime prevention effects of CCTV: A systematic review*. Technical Research Report No. 252, Home Office Research, Development and Statistics Directorate, UK.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: Harper Collins.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159-177.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*(3), 449-455.

Wicken, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance (3rd ed.)*. Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., & McCarley, J. D. (2008). *Applied attention theory*. London: CRC Press.

Wiedenbeck, S. (1999). The use of icons and labels in an end user application program: An empirical study of learning and retention. *Behaviour & Information Technology*, *18(*2), 68-82.

Wiener, E. L. (1973). Adaptive measurement of vigilance decrement. *Ergonomics, 16*(4), 353-363.

Wiener, E. L. (1987). Application of vigilance research: Rare, medium, or well done? *Human Factors*, *29*(6), 725-736.

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology, 25*(2), 229-235.

Wilkinson, R. L., Cary, J. W., Barr, N. F., & Reynolds, J. (1997). Comprehension of pesticide safety information: Effects of pictorial and textual warnings. *International Journal of Pest Management*, *43*(3), 239-245.

Wilkinson, R.T. (1964). Artificial 'signals' as an aid to an inspection task. *Ergonomics*, *7*, 63-72.

Wogalter, M. S., & Laughery, K. R. (1996). WARNING! Sign and label effectiveness. *Current Directions in Psychological Science*, *5*(2), 33-37.

Wogalter, M. S., Sojourner, R. J., & Brelsford, R. J. (1997). Comprehension and retention of safety pictorials. *Ergonomics*, *40*(5), 531-542.

Wogalter, M. S., Godfrey, S. S., Fontenelle, G. A., Desaulniers, D. R., Rothstein, P. R., & Laughery, K. R. (1987). Effectiveness of Warnings. *Human Factors*, *29*(5), 599-612.

Wolfe, J.M. (1994). Guided search 2.0: A revised model of search, *Psychonomic Bulletin & Review, 1,* 202-238.

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

Wolfe, J., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature*, *435*(7041), 439-440.

Wright, N., & McGown, A. (2001). Vigilance on the civil flight deck: Incidence of sleepiness and sleep during long-haul flights and associated changes in physiological parameters. *Ergonomics, 44*(1), 82-106.

Wright, M. C., Taekman, J. M., & Endsley, M. R. (2004). Objective measures of situation awareness in a simulated medical environment. *Quality and Safety in Health Care*, *13*, i65-i71. Retrieved from http://qshc.bmj.com/cgi/content/full/13/suppl_1/i65.

Yantis, S. (1993). Stimulus driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 676-681.

Yantis. S. (1996). Attentional capture in vision. In A.F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.). *Converging operations in the study of visual selective attention* (pp 45-76), Washington, DC: American Psychological Association.

Yantis, S. (2000). Goal directed and stimulus driven determinants of attentional control. In S. Monsell & J. Driver (Eds.), *Attention and Performance (Vol 18).* Cambridge: MIT Press.

Yantis, S., & Egeth, H. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 661-676.

Yantis, S., & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: Evidence from equivalent visual objects. *Journal of Experimental Psychology: Human Perception & Performance*, *20*, 95-107.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 601- 621.

Yantis. S. (1996). Attentional capture in vision. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds*.). Converging operations in the study of visual selective attention* (pp. 45-76). Washington, DC: American Psychological Association.

Yin, H. S. H., Dreyer, B. P., van Schaick, L., Foltin, G. L., Dinglas, C., & Mendelsohn, A. L. (2008). Randomized controlled trial of a pictogram-based intervention to reduce liquid medication dosing errors and improve adherence among caregivers of young children. *Archives of Pediatrics & Adolescent Medicine, 162*(9), 814-822.

van Voorthuijsen, G., van Hoof, H., Roubik, K., Klima, M. & Bernas, M. (2005). *CCTV Effectiveness Study*. TNO Defence, Security and Safety, The Hague. Unpublished Report. Also contained in paper presented at 39th International Carnahan on Security Technology. Las Palmas de Gran Canaria City, 2005.

Wallace, E., & Diffley, C. (1998). *CCTV Making it Work*, Control Room Ergonomics. HOSDB Publication 14/98.

Wallace, E., Diffley, C., Baines, E. & Aldridge, J. (n.d.) *Ergonomic Design Considerations for Public Area CCTV Safety and Security Applications*. Home Office, Police Scientific Development Branch (PSDB), UK.

Appendix 1: Access request letter

Psychology
School of Human & Community Development

University of the Witwatersrand
Private Bag 3, WITS, 2050

Tel: (011) 717 4500 Fax: (011) 717 4559

---

8 November 2010

Good day

CCTV RESEARCH.

My name is Fiona Donald and I am conducting research on ways of increasing CCTV surveillance operators' interest and attention. This is being done for studies towards a PhD degree at the University of the Witwatersrand. I am focusing on the operator's observation process and specifically looking at the effects of electronically inserting graphics into the video stream to enhance concentration and detection. Participants will be involved in viewing 1½ hours of video and will need to detect specified behaviours in order to evaluate their performance. For some participants, graphics of different kinds will be inserted to see which are most effective in improving the detection of incidents. I would like to invite your organisation and CCTV surveillance operators to participate in the study.

The research will occur in two stages on separate occasions and can be done on your premises. The first stage is to complete two observation exercises called SAMAE in order to benchmark their observation skills. Operators will be asked to provide consent for this process and it will take a total time of two hours. These assessments will take place free of charge and results will be made available to you if operators provide written consent for this. Alternatively, an outline of how the general results of your personnel compare to other organisations involved in the study can be provided.

The second stage of the research involves observing a video and responding to designated behaviours. This would occur on a separate occasion from SAMAE. Operators would watch the video individually and a research assistant or I would sit in and record responses. Briefing operators and watching the video will take about 1 hour 50 minutes in total per person. Operators will be shown how to respond to the video and will be given a short practice session immediately before watching it. This would be arranged at a convenient time for them.

No involvement would be required from you apart from setting up the sessions and providing a room where I can do the research. I would make all necessary arrangements, set up and administer the various exercises. I will provide a handout to operator personnel which can be distributed prior to the sessions to inform them about what would happen. Their signature on this form will be considered consent to participate in the study.

Participation by operators needs to be voluntary to ensure university ethics requirements are met, and employees will be assured that they will not be advantaged or disadvantaged in any way for choosing to participate or not to participate in the study. Because I will need to link their SAMAE and video observation results together, I will need to record their names. However, responses will be kept confidential and no information that could identify individual operators would be included in the research report.

Thank you for taking the time to consider taking part in this study and I would be happy to come through and discuss it further. I can be contacted telephonically at 083 6040 600 or via e-mail at Fiona.Donald@wits.ac.za. This research will contribute to a larger body of knowledge about vigilance, as well as to ways in which CCTV surveillance operators can be assisted in enhancing their detection of incidents. It also allows us to position South Africa as one of the leaders in this area. Feedback in the form of an executive summary will be provided to the organization and participants on request once the research is completed. A copy of the PhD research thesis would also be provided to the participating organisations on a complimentary basis.

Yours sincerely

Fiona Donald

Appendix 2: Participant information letter

School of Human & Community Development

Psychology
School of Human & Community Development

University of the Witwatersrand
Private Bag 3, WITS, 2050

Tel: (011) 717 4500   Fax: (011) 717 4559

---

8 November 2010

Good day

My name is Fiona Donald and I am a lecturer at the University of the Witwatersrand. I am doing research as part of a university degree. The research is on ways to make CCTV surveillance operators' jobs more interesting and easier. I am looking at how different types of pictures can help observation. Because you are an operator, I would like to invite you to be part of the study.

Below are some questions you may want to ask about the research.

1. **If I take part, what do I have to do?**
   If you decide to take part, you will need to attend two sessions:

   Session 1: You would do 2 observation exercises, called SAMAE and ScreenX. This takes 2 hours.
   Session 2: You would watch a video and tell us when you see certain behaviours. We would tell you beforehand which behaviours to look for. This takes another 2 hours.

2. **When will this happen?**
   We will arrange the sessions through your company at a time that suits you. These will be in November and December.

3. **Can I decide if I want to take part?**
   Yes, you can decide. Whatever you decide, it will not affect your job at all.

4. **Who will be told the results of the research?**
   For **Session 1**, this depends on what permission you give us. The results of the observation exercises will only be given to your company if you give your permission. You give your permission by signing Consent Form B (attached). If you decide not to participate or give your permission, that is ok. Your own results for **Session 2** will not be given to anyone and your name will not be mentioned in the report. A summary of the report will be given to your company, but no operators' names will be in this report either. Only results for all the operators together as a group will be in the report.

5. **What must I do next?**
   If you would like to be part of the research, please fill in both forms (attached). Consent Form A tells me whether you would like to take part. Consent Form B tells me if you give your permission for the results of the observation exercises to go to your company or not.

If you want to talk to me about the research, please phone me on [phone number]. Thank you for reading this.

Yours sincerely

Fiona Donald

Appendix 3: Consent forms

Consent Form A: Participation in research

I ____ consent to participate in the study by Fiona Donald on vigilance in CCTV surveillance operators.

I understand that:

- Participation is voluntary
- My name and personal details will not be included in the research reports
- My responses are confidential
- Individual results of the video observation exercise will not be given to management.
- I will not be advantaged or disadvantaged in any way by participating in the study
- I may withdraw from the study at any stage while watching the video if I wish to do so.

I understand that data will be stored securely by Fiona Donald on completion of the study.

Name: ___________________________________ Signed: _____________________

Phone number (w): _____________________ Cell: _____________________

Company:____

Consent Form B: SAMAE

Please cross out the words that do not apply to you:

- I do / do not give consent to complete the SAMAE and ScreenX exercises
- I do / do not give consent for the researcher to give my SAMAE and ScreenX results to the organization I work for.
- I understand that data will be stored securely by Fiona Donald on completion of the study.

Name: ___________________________________ Signed: _____________________

Phone number (w): _______________________ Cell: _______________________

Company name: ____

Appendix 4: Log sheet example

Participant:				**Group:**	**Research assistant:**		
Clip start time	**Area**	**Activity or nature of TB**	**Time of TB/IGO**	**Response required**	**Response given: Time and behaviour**	**Score: Hit, FA, correct rejection**	**Non-verbal behaviour: Time & description**
0h00m00s	CDX	Sweeping					
0h01m10s	CDX	Middle man takes diamond off side of machine	01m37s	Pick up			
0h01m45s	X-ray	Enter, X-rayed, exit					
0h02m17s	Large machine	Hosing machine, watching					
0h04m12s	CDX	Sweeping spillage					

Appendix 5: Sequence and times of target behaviours and IGOs.

Time (hh:mm:ss)	Time interval (mm:ss)	Target behaviour	IGO
00:01:37	01:37	Pick up	
00:06:34	04:57	Pick up	
00:08:30	01:56		IGO
00:10:37	02:07	Swivel	
00:12:00	01:23		IGO
00:16:05	04:05		IGO
00:17:59	01:54	Broom flick	
00:23:12	05:13	Pick up	
00:26:31	03:09	Kick	
00:28:30	01:59		IGO
00:31:09	02:39	Pick up	
00:33:59	02:40	Swivel	
00:37:00	03:01		IGO
00:40:20	03:20		IGO
00:42:14	01:54	Pick up	
00:45:30	03:16	Pick up	
00:50:31	05:01	Pick up	
00:54:00	03:29		IGO
00:57:18	03:18	Broom flick	
00:59:30	02:12		IGO
01:00:53	01:23	Broom flick	
01:03:30	01:37		IGO
01:06:07	02:37	Kick	
01:09:31	03:24		IGO
01:11:30	01:59		IGO
01:12:42	01:12	Pick up	
01:14:39	01:57	Kick	
01:21:42	07:03	Pick up	
01:25:20	03:48		IGO
01:28:15	02:55	Pick up	

Appendix 6: Ethics approval certificate

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (NON MEDICAL)
R14/49 Donald

CLEARANCE CERTIFICATE **PROTOCOL NUMBER H0 90625**

PROJECT Optimal characteristics of inserted graphics objects in stimulating CCTV operator vigilance and performance

INVESTIGATORS Mrs F Donald

DEPARTMENT Psychology

DATE CONSIDERED 12.06.2009

DECISION OF THE COMMITTEE* Approved Unconditionally

NOTE:

Unless otherwise specified this ethical clearance is valid for 2 years and may be renewed upon application

DATE 20.07.2009 **CHAIRPERSON** ....
(Professor R Thornton)

cc: Supervisor : Prof Y Broom

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10005, 10th Floor, Senate House, University.

I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

____
Signature

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix 7: TB hits and misses by participant

	Target behaviour																	
	Phase A						Phase B						Phase C					
Obs	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
1	X	X				X	X	X	X							X		
2	X	X	X		X	X		X	X		X		X	X	X	X	X	
3		X	X	X		X	X	X	X			X	X	X		X	X	
4	X	X		X		X		X	X					X	X	X	X	X
5	X	X		X		X			X	X	X		X	X	X	X	X	X
6	X	X			X	X	X	X	X		X				X		X	
7		X	X					X									X	
8	X	X		X		X					X		X		X			
9	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X
10	X	X	X			X		X	X	X	X		X	X	X	X	X	X
11						X	X	X	X	X	X		X		X		X	
12				X	X			X	X				X		X			
13				X			X	X					X	X				
14	X	X		X				X							X	X	X	X
15	X	X	X	X	X	X	X	X	X		X			X		X		
16	X	X	X		X			X	X						X	X		X
17		X			X	X			X		X		X		X			
18	X		X			X		X	X				X					X
19	X	X	X			X		X	X		X			X	X		X	X
20	X			X	X	X	X	X	X			X	X		X		X	
21	X				X		X	X	X		X		X	X	X	X	X	X
22	X	X	X	X		X	X	X	X		X			X	X	X	X	
23	X	X	X	X	X	X	X	X	X	X		X	X		X			X
24	X	X		X		X	X	X	X						X	X	X	X
25	X		X	X		X		X	X			X	X	X	X	X	X	X
26	X	X		X		X		X	X				X			X	X	
27	X							X					X	X	X	X		
28					X	X		X	X				X		X	X		
29			X				X	X	X				X		X			
30	X	X		X		X		X	X		X		X	X		X	X	X
31	X	X			X			X	X		X		X		X		X	
32	X	X	X	X	X	X	X	X	X		X		X	X	X	X	X	X
33	X		X	X	X		X	X	X		X		X		X		X	X
34		X		X	X	X	X	X	X		X		X	X	X	X	X	X
35	X	X					X	X	X						X	X	X	X
36	X	X	X	X		X		X	X		X		X	X	X		X	
37		X	X			X			X		X			X	X		X	X
38	X	X	X			X	X	X	X		X		X	X	X		X	X
39			X	X	X	X	X	X					X		X			
40	X			X					X						X			X

Appendix 7: TB hits and misses by participant cont.

	Target behaviour																	
	Phase A						Phase B						Phase C					
Obs	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
41		X						X						X	X			
42		X	X	X		X		X			X				X	X		
43	X		X			X	X	X	X		X		X		X	X	X	
44	X	X	X	X			X	X	X	X								X
45	X	X			X			X	X					X	X	X	X	X
46	X		X								X							
47	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	
48	X	X	X			X		X	X					X	X	X	X	
49	X	X			X	X	X	X			X							
50			X			X		X					X	X			X	
51	X	X	X	X		X										X		
52		X	X		X	X		X	X					X		X		
53			X			X		X					X	X	X	X		
54	X	X	X	X		X	X	X	X						X			X
55									X							X	X	
56		X				X		X	X				X		X	X	X	X
57	X	X	X			X		X	X							X		
58						X			X						X			
59	X		X			X							X			X		
60			X			X								X				
61	X	X			X	X	X	X	X		X					X		X
62	X	X	X		X		X	X							X	X	X	
63	X			X		X			X					X		X		
64	X	X	X	X		X		X	X	X			X		X	X	X	X
65			X		X	X		X	X	X					X	X		
66					X	X		X								X		X
67	X	X	X	X		X	X	X	X							X		
68	X				X			X	X						X		X	
69	X	X	X			X	X		X				X	0	X	X	X	X
70	X	X	X	X	X	X	X	X	X	X		X	X		X		X	X
71	X						0	X								X		
72						X		X							X	X		
73				X		X	X		X						X		X	X

Note: X denotes a 'hit.'