

## **Applying psychological science to the CCTV review process: A review of cognitive and ergonomic literature.**

Anne P. Hillstrom, Ph.D.

Lorraine Hope, Ph.D.,

and

Claire Nee, Ph.D.

University of Portsmouth

Department of Psychology

21 March, 2008

This report was commissioned by and delivered to the Home Office Scientific Development Branch.

British Crown Copyright Reserved 2008. Reproduced with the permission of the Controller of Her Majesty's Stationary Office.

## Table of Contents

Executive summary	ii
Recommendations	iii
1. Introduction	1
2. The task	2
3. The Degraded Nature of CCTV Footage	4
3.1. Object Recognition	4
3.2. Face Recognition	6
3.3. Event Recognition	12
4. Distractions and Interruptions	17
4.1. Distractions	17
4.2. Interruptions	18
5. Time on Task	20
5.1. Maximizing comfort over long periods of time spent reviewing footage	20
5.2. Maximizing performance over long periods of time spent reviewing footage	21
6. Mindset	28
6.1. Maintaining the ability to detect unexpected things	28
6.2. Mindset in making decisions	31
7. Variability in people	39
7.1. Personality and other individual differences that affect performance	39
7.2. Training	43
8. Report conclusion	48
References	49

## ***Executive Summary***

As CCTV cameras are used more and more often to increase security in communities, police are spending a larger proportion of their resources, including time, in processing CCTV images when investigating crimes that have occurred (Levesley & Martin, 2005; Nichols, 2001). As with all tasks, there are ways to approach this task that will facilitate performance and other approaches that will degrade performance, either by increasing errors or by unnecessarily prolonging the process. A clearer understanding of psychological factors influencing the effectiveness of footage review will facilitate future training in best practice with respect to the review of CCTV footage. The goal of this report is to provide such understanding by reviewing research on footage review, research on related tasks that require similar skills, and experimental laboratory research about the cognitive skills underpinning the task. The report is organised to address five challenges to effectiveness of CCTV review: the effects of the degraded nature of CCTV footage, distractions and interrupts, the length of the task, inappropriate mindset, and variability in people's abilities and experience. Recommendations for optimising CCTV footage review include (1) doing a cognitive task analysis to increase understanding of the ways in which performance might be limited, (2) exploiting technology advances to maximise the perceptual quality of the footage (3) training people to improve the flexibility of their mindset as they perceive and interpret the images seen, (4) monitoring performance either on an ongoing basis, by using psychophysiological measures of alertness, or periodically, by testing screeners' ability to find evidence in footage developed for such testing, and (5) evaluating the relevance of possible selection tests to screen effective from ineffective screeners.

## ***Recommendations***

The literature review that follows has resulted in the following recommendations for how to optimise CCTV footage review.

- CCTV review is a cognitive task yet the cognitive skills involved in successful CCTV review are both manifold and difficult to describe. For that reason, we recommend a Cognitive Task Analysis (CTA) of the process to identify particular areas of difficulty and enhance future performance. CTA is critical for cognitive training and “to improve cognitive functions such as decision making, sense making, problem detection, replanning” (Crandall, Klein & Hoffman, 2006, p.214). A Cognitive Task Analysis will also allow the production of meaningful scenarios useful for training purposes.
- Encourage the exploitation of new technology so that use of degraded footage is minimised. Footage that is degraded by visual noise, poor image resolution, or minimal frame rates will degrade identification, although other factors may be just as detrimental and wash out the effect of perceptual difficulties. In terms of technology currently available, the use of colour footage as opposed to B/W appears to have no benefit to reviewers but recording colour footage requires so much more computer memory that images are often compressed more in frame rate or spatial resolution. Therefore, the use of B/W footage should be encouraged.
- Brief reviewers fully on the investigation before asking them to review footage. Having erroneous expectations about what they are looking for will affect negatively reviewers’ performance.
- Train reviewers about change blindness and about typical biases in decision making. Those who are taught about these threats to performance are less susceptible to them. Although one would expect that good investigators might already have cognitive flexibility, some people might benefit from scenario-based experience in taking alternative perspectives to avoid too rigid expectations in how to view and use footage.
- Have reviewers be vigilant to make sure target in one footage is the same in other footage. People are not good at detecting differences and frequently believe they will be better than they are.
- Reviewing should be done in a room where the reviewer is alone, to avoid interruptions and distractions.
- For the reviewer’s health and safety and to minimise vigilance decrements, follow ergonomic standards on computer workstations, optimise lighting in the room, provide

temperature control, even if it is by means of a window or a fan, and provide a radio to allow people to vary the noise environment to aid alertness.

- Reviewers should be encouraged or required to take short breaks regularly to maintain vigilance and to avoid the muscle strain that can result from sitting too long. The breaks should be initiated by reviewers rather than others, so that the reviewer can manage the interruption to the task optimally.
- Reviewers will do a better job if they are not sleepy to start with.
- Perceptual training for staff who are new to reviewing CCTV footage might start with having people look for specific people, targets, or actions in a relatively simple scene context. Using simple rather than complex scenes will maximise what is learned about the appropriate cues to recognition in degraded images.
  - Ensure training images include a range of types of video degradation so that people learn to make difficult identifications as well as simple ones.
  - Instructions should highlight the features that can be used to aide identification.
- Train vigilance after training recognition. To train vigilance, give practice in which there is the potential for people to miss targets, and give feedback when they do. Use the same kind of searches that would be used on the job.
- To train use of footage, give knowledge, build rules, and develop automatic prioritisation of underlying skills.
- The following skills and/or individual differences could be considered when setting up selection criteria for reviewers. All of these require research validation to be sure they are relevant to this particular task:
  - Knowledge of the kinds of criminal behaviour that needs to be identified.
  - Ability to see perceptual differences (visual tests and perception tests)
  - Ability to recognise faces (Cambridge Face Memory Test).
  - Ability to detect targets across a broad spatial span when looking at one location in an image (Functional Field of View).
  - Ability to see camouflaged objects (hidden pictures tests)
  - Ability to maintain focus amid distractions (distractability test, tests of executive control)
  - Ability to react to information that fails to match expectations (cognitive flexibility tests)
  - Ability to maintain vigilance (Psychomotor Vigilance Test)

## ***1 Introduction***

As CCTV cameras are used more and more often to increase security in communities, police are spending a larger proportion of their resources, including time, in processing CCTV images when investigating crimes that have occurred (Levesley & Martin, 2005; Nichols, 2001). As with all tasks, there are ways to approach this task that will facilitate performance and other approaches that will degrade performance, either by increasing errors or by unnecessarily prolonging the process.

The goal of this report is to consider the cognitive processes involved in the task of reviewing CCTV footage. A clearer understanding of the task at a psychological level will facilitate future training in best practice with respect to the review of CCTV footage. The report is organised as follows. First, we describe our understanding of the task of reviewing CCTV footage. Then we review psychological literature that addresses the major challenges to effective CCTV review: the effects of the degraded nature of CCTV footage, distractions and interrupts, the length of the task, inappropriate mindset, and variability in people's abilities and experience.

To write this report we drew first on published research about CCTV usage. However, there has been very little research about human performance in reviewing CCTV footage, so we also drew heavily on research about other workplace tasks that require somewhat similar inspection and search skills. To deepen understanding of the findings, we drew on experimental laboratory research about the underlying cognitive skills involved.

## **2 The task**

It is important to establish a common ground of understanding of the nature of the task. It would be useful to conduct a formal task analysis, as has been done for many other workplace tasks. But even an informal exploration suggests that the task is extremely varied. To start with, the reviewer could be searching for a person, an object, an event, or interpreting what is seen to establish a timeline of complex events.

The task is generally *long*. It is reasonable to assume that most times police need to review CCTV footage, they are spending hours, not minutes, on the task. A review may sometimes require only minutes, as when looking for a piece of evidence to support the case for a crime that is thoroughly understood, and other times may require hundreds of hours, as when trying to lay out the timeline of a very complex and serious crime, such as terrorism.

The task *requires attention*. Footage today is often degraded in terms of image resolution (not enough pixels making up the image), colour (distorted away from what is natural) and motion resolution (not enough frames per second showing action). What is being searched for can be difficult to see and may appear only briefly. The visual appearance of the target can be uncertain and the time of appearance can be uncertain.

The task *requires elaborative and interpretive abilities*. Sometimes a concrete, known target is being sought, but even then its appearance may be different than expected, as when a target person disguises himself or herself. Other times something specific is being sought but is not fully defined in terms of its perceptual features. For instance, when looking for a suitcase, the appearance could be quite different depending on its type or brand. Other times the reviewer is trying to understand a sequence of events, and so must integrate information over time and space (e.g., events recorded from multiple cameras). In addition, when looking at even one very small window into the world, one must use expectations about what should or could be occurring to understand the snippets of behaviour that are seen.

*Expectation* plays a key role in the task. Most of the time, the officer conducting the review will be well briefed on the case before searching through footage. However, unlike someone watching live surveillance, a person reviewing footage knows that an event has taken place, and so has an expectation, albeit not always accurate, of what they are looking for.

The task is *not engaging*. Reviewing footage is one of many tasks an officer could be doing and seldom is the preferred task, in terms of a match between what police officers find interesting to do and what the task requires them to do. Further, events recorded are seldom intrinsically interesting for anyone to watch.

The task can have *important outcomes*. Establishing good evidence on CCTV footage can lead a suspect to confess, saving much time, money and resources in an investigation (Levesley & Martin, 2005). Even if it does not lead to a confession, CCTV evidence can aid swift convictions if a case goes to trial (Costigan, 2007).



### ***3 The degraded nature of CCTV footage***

#### **3.1 Object recognition**

Human recognition of objects is so effective in general that we tend to be unaware of how much variability our visual system is able to ignore. The ability to understand the gist of what we are seeing depends on seeing the three dimensional parts of the object first, along with their spatial relationship to one another, and comparing what we see to our memory (Biederman, 1987). This processing is relatively insensitive to the viewpoint from which we are viewing the object compared to how we have seen it before. For finer discriminations, such as discriminating our car from another person's car, our recognition is more sensitive to the viewpoints we have experienced (Peissig & Tarr, 2007) and to the outline of the objects (Stone, Buckley & Moger, 2000). Even so, our visual system is able to ignore all but the most extreme differences in lighting and shadow, which can render invisible some of the object's contours and even change the object's colour, and the fact that we often are not seeing the entire shape of what we are trying to recognise, since objects are often partly blocked by other objects as we look at them.

There have been a number of studies that investigated the effect of different kinds of image degradation – low luminance contrast, low frame rates in videos, and poor spatial resolution on recognition. As we present this section, we will include some technical details that are perhaps too detailed for some readers of this report but that may aid those who need to make decisions about the limits of acceptable image quality.

**Contrast.** Luminance contrast can be thought of as the amount of difference between light and dark shades in an image. Snyder (1974) presented a measure of television image quality based on luminance contrast ratios available, and showed that image quality reliably predicts the ability of people to match TV-based large images of people's faces when compared to sharp photographs. A later study showed that luminance contrast also affects reading speed on both cathode ray tubes (CRTs), the most common kind of monitors until very recently, and on paper (Snyder & Jorna, 1991).

Leermakers and Boschman (1984) and Boschman and Roufs (1989; 1997) found that luminance contrast is a strong determinant of performance when searching for letters in pseudo-text (strings of letters with spaces interspersed occasionally), reaching asymptote at about 0.4 log contrast ratio, and that when contrast was held constant, the sharpness of the image (due to bandwidth cutoff) impacted objective performance up to about 6 cycles/degree

and impacted subjective ratings of visual comfort beyond that (see also Näsänen, Karlsson & Ojanpää, 2001).

Näsänen and Ojanpää (2003) asked participants to search for icons on a typically coloured computer desktop and varied both luminance contrast and sharpness (contrast in only the high spatial frequencies, more commonly described as blur). They found that contrast affected search time when it was low, but beyond about 0.4 search times grew no faster. Sharpness had a small effect, and up to approximately 5 cycles/degree. The number and duration of eye fixations, too, were affected more by contrast than by sharpness.

**Frame rate.** The screen refresh rate of a monitor affects eye movements when observers are reading (Kennedy & Murray, 1991) and when they are instructed to look at briefly presented targets in a display of words (Kennedy, Brysbaert, & Murray, 1998). This occurs even when the refresh rate is high enough not to produce visible flicker, and is hypothesised to occur because the pulsing luminance that does not reach awareness activates the part of the visual system that processes transient stimuli, which influences eye-movement control. A video running with a low frame rate would be expected to have much the same effect, as it would produce a similar pulsing of background luminance changes in addition to the informative luminance changes caused by the dynamic events shown. In keeping with this, a study asking people to search through static shapes superimposed by random moving noise (as used to be common on television screens) found that the moving noise was more disruptive to search accuracy and search time when it was slow (5.2 fps) than when it was fast (26 fps) (Erickson, 1966).

**Spatial resolution/Noise/Sharpness.** Mocharnuk, Gaudio, and Suwe (1981) simulated the images produced on imaging infrared (IIR) systems used to detect and identify ships when at sea. Distance to the target always appeared to be getting smaller. They varied image quality, frame rate, luminance contrast ratio in the display, and system resolution. Frame rate had no effect, perhaps because, although the dynamics in the display were instrumental to achieving identification, they did not provide information about the object. Luminance contrast improved identification, in the sense that identification could be achieved at further distances when contrast was greater. When images were degraded by imposing bandpass filters, which would limit the spatial resolution of the images, more degraded images were identified less accurately at each distance.

Pearson and Pearson (1985) used static aerial reconnaissance photos transmitted over a communication line, and simultaneously varied original image quality and bandwidth

compression in a study of ability to make a variety of judgments of the pictures. The ability to count targets, which requires only object segregation, the ability to fully identify rather than only coarsely categorise the object, and the speed of judgments were all affected by image quality and compression ratios. Quality and compression interacted for the completeness of identification and speed of identification, but not for the accuracy of counts.

**Summary and Implications for CCTV footage review.** It will be more difficult to recognise objects when luminance contrast is low and when spatial resolution is coarse. Slow frame rates may interfere with eye movements to scan static pictures of objects, which suggests interference with object recognition. There will be a point at which further improvements in these factors will not show much more improvement, but it is difficult to quantify what that point will be, as it will depend on how small the objects to identify will appear in footage. It is our understanding that image quality of footage available for review (not necessarily the original footage) is currently of quite poor quality. It is important, then, to exploit any advances in technology available to improve image quality.

There is currently quite a lot of research being done by computer scientists and engineers on automation of object recognition, face recognition, and action recognition in CCTV footage (Hesse, 2002). Although the primary use of such automation is for reducing the workload of monitoring *online* images, it may be useful in the future for offline footage review. We have not reviewed the developments in automation, as it generally is not psychological research. However, there are psychological factors to its effectiveness, as automation is almost always controlled by people and used to aid people's decisions (e.g., about decisions about whether to send an officer to intervene in an event that has been detected). One important psychological factor is the degree of trust the person places in it. The effectiveness of any automation aid to human decision making is strongly influenced by the degree to which the decision-maker relies on it. When the automation is imperfect, people tend to misjudge the degree to which they should rely on it versus doing their own search of the evidence. In fact, a review of the literature suggests that an aid that makes correct decisions less than 70% of the time is worse than having no aid at all, in terms of how it impacts the performance of the person using the aid (Wickens & Dixon, 2007).

### 3.2 Face Recognition

Faces are among the most subtly different "objects" we have to discriminate in everyday life. The geometry of faces is overall the same, and the differences between some faces are very minor. In fact, face recognition usually is based on holistic processing rather than a coarse

determination of the three-dimensional parts that constitute a face and their spatial relationship to one another (Tanaka & Farah, 1993). Although people are very used to identifying people based on their facial appearance, memory for faces of people who are not well known (i.e., faces of people we have only recently encountered) is more sensitive to the less permanent aspects of faces, such as hair style or expression on the face (e.g., Wright & Sladden, 2003). Movement also plays a role in recognizing faces. (Hill & Johnston, 2001; Roark, O'Toole, Abdi & Barrett, 2006). Even when we cannot see a person's face, the movements of the person will often help us to recognise them visually (Jastorff, Kourtzi & Giese, 2006; Jokisch, Daum & Troje, 2006; Roark et al, 2006). Movement also aids recognition of objects. For instance, when a novel object was shown during familiarization moving in a certain pattern, but then recognition was tested showing the movie running in reverse, recognition slowed substantially (Stone, 1999; see also Kourtzi & Nakayama, 2002).

**Spatial resolution.** Using a bandpass filter to limit whether the fine or coarse details of a picture are available has shown that face recognition is more impacted by loss of fine detail than loss of coarse detail (Costen, Parker & Craw, 1996; Fiorentini, Maffei & Sandini, 1983). The results are reported in terms of the amount of detail needed given a certain face width, which of course will vary according to how far away the face appears. So the required spatial resolution of the image depends on the size of the face in the image, a situation which is already captured in the guidelines the Home Office gives to people who are setting up CCTV cameras (Cohen, Gattuso & MacLennan-Brown, 2006). In fact, observers who are trying to recognise people at a far distance tend to rely on external features, such as clothing, more than faces (e.g., Tickner & Poulton, 1975).

### **Face recognition in the forensic context**

The unreliability of eyewitness identifications where witnesses have to rely on their memory of a face is well documented (Wells & Olson, 2003). Poor eyewitness identification accuracy has been observed in both laboratory and archival studies. For instance, in their analyses of police archival data Wright and McDaid (1996) noted that the choice of foil (or known stand-in) was about 20%. It is generally acknowledged that memory for faces can be error-prone particularly under certain conditions, such as short exposure duration (Memon, Hope & Bull, 1993), when the perpetrator is disguised or when a weapon is involved (Steblay, 1992).

Intuitively, one might expect that identification performance might improve significantly when the 'witness', be that the original witness, a CCTV operator or police officer reviewing the evidence, has access to a video recording of the (alleged) target and, possibly, still

photographs of the suspect. A witness would no longer need to rely on their memory or, in the case of a police officer, prior familiarity with the suspect. With video footage of the incident available, the task would no longer rely so heavily on memory and would simply require the witness to engage in an apparently simple matching task – particularly where still photographs were also available. Put simply:

“you may think there is all the difference in the world between a fleeting encounter with one’s eyes over a second or two and a permanent record on video type tape which is there to be played over and over again” (Extract from the trial judge’s direction to the jury; cited in Costigan, 2007)

However, research by psychologists has revealed that the identification of individuals from CCTV footage is not necessarily a simple identification task and, like other identification tasks, is prone to error – even under optimal conditions.

In this section, the difficulties associated with identifying suspects from CCTV footage will be examined in light of research evidence. In particular this review will highlight the difficulties that might be encountered when establishing the identity of a suspect on CCTV. There are two quite distinct circumstances where an attempt may be made to identify a face from a video image (Bruce et al., 1999). In the first situation, a spontaneous identification may be made by a member of the public (or perhaps, a CCTV operator or police officer) who claims that the target appearing in the CCTV image is personally known to them. In the second situation, the target appearing in the CCTV footage is compared to an apprehended suspect to establish whether, in fact, the suspect was recorded at the scene of the incident under investigation. Research suggests that identification accuracy varies under these circumstances with respect to whether the face is previously known or previously unknown to the witness.

**Previously unknown faces.** Early research on face matching ability adopted Benton’s neuropsychological tool (the Benton Face Recognition Test (Benton, 1980)). In this test which is typically used to measure particular forms of neuropsychological dysfunction, participants are required to identify which face from a set of faces matches a particular target face. At the outset, the target face is matched to an identical photograph from a set of six alternatives. As the task progresses, changes to the lighting and viewpoint make the task rather more challenging. When people who show no signs of dysfunction attempt the task, the error rate approaches 20%, which suggests that simply identifying and matching a target face can pose difficulties even without a memorial component. This phenomenon has also

been documented in laboratory-based studies of unfamiliar face matching (Hill & Bruce, 1996).

In one of the early studies on spontaneous identifications based on prior exposure, Logie, Baddeley and Woodhead (1987) examined the ability of the general public to identify a live target in a town centre from a previously presented photograph. The photograph had been published in a local newspaper. Despite details on the precise location of the target, the spontaneous detection (i.e. identification) rate for the general public was very low and this was coupled with a high false recognition rate (i.e. false identifications of other 'innocent' passersby).

These low recognition rates in dynamic interactions where the target face is continually available to the witness have been documented elsewhere. Kemp, Towell and Pike (1987) conducted a field study to examine whether credit cards bearing a photograph of the cardholder might serve to reduce credit card fraud. Including a photograph of the legal cardholder on a credit card (or indeed, other identity document) would seem to be a relatively foolproof method of ensuring the card is only used by the person entitled to use it. In their study, shoppers presented a credit card bearing a photograph of themselves to pay for half the transactions while for other transactions they presented a card bearing the photograph of another individual. When the photograph was of someone other than the shopper, it sometimes depicted an image of a person judged to resemble the shopper in appearance (a 'matched' foil). For other transactions, the photo depicted a person judged to be dissimilar in appearance to the shopper appeared on the card. In all conditions, the photographs were of a uniformly high quality and were no more than 6 weeks old. Experienced checkout cashiers were required to either accept or decline the card depending on their verification of the cardholder's identity, and rate their confidence that the photograph appearing on the card was, in fact, that of the shopper. More than 50% of the fraudulent cards were accepted by the cashiers – despite the fact that cashiers were aware that a study was underway and acknowledged that they both spent longer examining cards and had been more cautious than usual. When the photograph resembled the shopper, only 36% of the cashiers correctly declined the card. Despite these high error rates, behavioural observations suggested that the cashiers had spent some time during the transactions deliberately comparing the appearance of the photograph and the shopper.

High error rates in the ability to match a target from CCTV footage have also been documented. Typically, it has been assumed that difficulties in identifying faces from video recordings are largely due to the frequently poor quality nature of the recording and that

were highly quality recordings available such difficulties would not arise. While it is true that many CCTV images may be of poor quality for a number of technical reasons (such as unsuitable lighting conditions, intermittent image sampling etc.), the assumption that this alone underpins low accuracy rates in face matching from CCTV has been challenged by research findings.

Bruce and her colleagues (1999) examined how well people were able to match faces extracted from a high-quality video-recording against high quality photographic images. The results revealed that overall accuracy was relatively poor (averaging only 70% across trials) even under these optimal conditions (i.e. identical photograph matching with high quality target images). Performance was further degraded when the target expression or viewpoint was altered. Furthermore, the use of colour target images (as opposed to black-and-white images) did not appear to lend any particular advantage (or disadvantage) to performance on the matching task. Henderson, Bruce and Burton (2001) extended these findings and in a series of five studies found that the ability to match a target appearing CCTV footage to another image of that same target was highly error-prone. Similarly, Davies and Thasen (2000) conducted two studies examining matching ability from both face and whole body CCTV images. In the first study, they examined identification ability from CCTV recordings taken in a large public space (a car park). Accuracy rates for matching were in the region of 30% despite the fact that participants had the opportunity to consult a constant still frame of the target. Davies and Thasen (2000) explain this low accuracy rate with reference to the change in perspective between the CCTV image and the photograph of the target. The CCTV images of the target were recorded from a height of six metres above ground level while the photographs were taken at eye-level. In the second study, the CCTV images were taken at close range to the target (such as might be produced by a surveillance camera in a bank or pay-point). Thus, participants saw high-quality full-face colour images of the target. However, matching results remained poor – despite the optimal conditions, only 56% of participants correctly matched the CCTV image to the target photograph. The research results reviewed in this section indicate – with some consistency – that our ability to identify an unfamiliar face – even in the presence of a reference image (such as a still or a photograph) is surprisingly error-prone.

**Known or Familiar faces.** In contrast, identification accuracy for known or familiar faces can be very accurate – even when the target images are of poor quality. To examine the impact of familiarity on face recognition, Burton, Wilson, Cowan and Bruce (1999) showed study participants surveillance video footage of a target (a lecturer) who would be known to some participants but not others. Three groups of participants were recruited. Two of the groups comprised students who were either familiar or unfamiliar with the target. The third group

comprised police officers who were unfamiliar with the target but who were experienced in making identification judgements (average of 13.5 years of service). After being shown the video footage, participants were asked to identify the target from a set of high quality photographs. Results indicated a marked advantage for people who were personally familiar with the target – 73% of the poor quality image targets were recognised *when they were familiar*. Regardless of whether they were students or police officers, people who were unfamiliar with the target performed very poorly on the identification task. In a second study, Burton et al. (1999) explored whether this familiarity effect was due to the recognition of factors such as target gait or body shape. Participants were shown video surveillance clips of a familiar target. The clips were edited such that the body, face or gait of the familiar target was obscured. Results suggested that the advantage for familiarity was largely due to recognition of the face, rather than the recognition of other cues such as gait and body shape as identification accuracy was significantly worse when the face was obscured. Bruce, Henderson, Newman and Burton (2001) extended this research and, in a series of studies exploring the role of familiarity, found that participants were able to correctly verify (or reject) a familiar target with a high degree of accuracy (over 90%) despite the use of poor quality video images. When participants were unfamiliar with the targets, the accuracy rate was significantly lower (56%). Subsequent experiments revealed that brief periods of exposure to the target do not necessarily generate sufficient familiarity to improve the recognition or matching of unfamiliar faces – unless some ‘deep’ or social processing has taken place (i.e. discussing the faces with another person).

**Summary and Implications for the review of CCTV footage.** The psychological research literature reviewed above clearly demonstrates that both the matching and identification of a previously unfamiliar face from CCTV footage is a surprisingly difficult and error-prone task. However, by contrast, accuracy rates for the identification of known faces – even from poor quality footage – are typically high. The research also demonstrates that colour footage does not lend any particular identification advantage. Finally, while there is some (modest) evidence to suggest there may be some naturally occurring individual differences in face matching ability, there is no evidence to suggest that people with experience of the identification process (such as experienced police officers) perform any better than untrained, inexperienced individuals when required to match or identify a previously unknown face.

Face recognition is of central importance to investigative police work (Scott-Brown & Cronin, 2007). CCTV has the benefit of providing investigators with a permanent record of an event and, importantly, who may have been involved in it. The availability of CCTV footage – and



the speed at which it was analysed – facilitated the rapid identification of the 7/7 and 21/7 bombers from thousands of hours of recordings (Metropolitan Police, 2005). Furthermore, actual CCTV footage is generally considered powerful evidence in court (NACRO, 2002; Scott-Brown & Cronin, 2007; Thomas, 1993).

However, relying on CCTV for the recognition and identification of suspects may foster a false sense of security and a potentially dangerous over-reliance on such evidence. Costigan (2007) emphasises the risk of injustice in a review of recent appeal cases involving disputed identifications from CCTV (e.g. *Brady; Clare and Peach; Dodson and Williams* – see Costigan, 2007 for an examination of the legal issues relating to CCTV evidence). Certainly, the research reviewed above clearly illustrates the problematic nature of unfamiliar face recognition. These findings are counter-intuitive to the lay assumption that identifying a face captured on CCTV by matching it to either a suspect photograph or the actual live suspect should be a simple, objective and non-fallible task. In other words, people expect to be able to do this task with a high degree of accuracy. However, the research consistently demonstrates that people are poor at this task – even under optimal conditions.

A further lay assumption reported among CCTV stakeholders, such as businesses (see Davies & Thasen, 2000) is that colour CCTV images will be more useful in the investigation of crime than black-and-white images. With respect to facial identification, several studies have shown no advantage for colour CCTV footage.

### **3.3 Event recognition**

Compared to recognition of objects and people, recognition of actions has only begun to be researched. Although a complex action involves continuously flowing motion, the actions are understood as involving distinct parts (Newtson, Enquist & Bois, 1977, as cited by Saylor & Baldwin, 2004). When people are asked to describe actions they see, they report a quite structured hierarchy of goal-directed movement, with larger actions made up of more detailed actions. Descriptions people give of the topmost level structure generally give most detail of the objects involved; descriptions given of the low levels give most detail of the actions involved (Zachs, Tversky & Iver, 2001). Further down the hierarchy, the detailed movements involved are described with less links to intention or objects. The organisation is partly learned through experience (Avrahami & Kareev, 1994), but since infants appear to see much the same organisation in action as adults do, there must be some basic features of movements that in themselves signal the start or end of parts of actions without extensive learning (Baldwin, Baird, Saylor & Clark, 2001). Current research is trying to identify those

features. In fact, the intention of human actors strongly affects how we describe their physical actions (Saylor & Baldwin, 2004), and so our understanding of the social context of what we are seeing influences event perception.

Event comprehension for complex and unfolding events is a lot like comprehension of text. Research about text comprehension, then, can be useful in characterising the cognitive processing involved (e.g., Kintsch, 1988). First, it is a sequential task, where information presented early must be remembered to interpret the significance of information presented later. This requires the observer to keep in mind enough detail that at some point conscious or unconscious decisions will need to be made about which information is most important to the ongoing “plot”. Second, comprehension requires evaluating what currently is being seen against not only earlier parts of the event, but also long term memories for related things (e.g., knowledge of the known people seen, how similar events encountered in the past unfolded). Failures in comprehension, then, can result either from missing a critical but briefly presented point, from misjudging what is critical and so forgetting critical information, or from failing to see the associations between current events and long past events.

**Frame rate.** Earlier we pointed out that, at least with static displays presented on video, a slow frame rate might disrupt eye movements, presumably due to the transient perceptual changes caused at the transition between frames (Erickson, 1966; Kennedy et al, 1998). However, potentially inconsistent results have been found in a study using dynamic videos that had frame rates from 5-25 fps in multimedia videos. The videos were of television shows and what was measured was the median gaze position for each video at each point in time for each group of participants (Gulliver & Ghinea, 20). When comparing participants watching the same videos at different frame rates, strong correlations were found between the positions gazed at in the video as the video progressed. Although this is not the same kind of measure of disruption to eye-movements, this measure’s lack of sensitivity to frame rate does suggest that when viewing dynamic displays, eye movements are not severely disrupted by large changes in frame rates.

A few studies have asked participants to rate videos on their acceptability while varying things like frames per second, signal to noise ratio and spatial resolution (e.g., Apteker et al, 1995; Cranley et al, 2003; Fukuda et al, 1997; Hikichi et al, 2001; Yamazaki, 2001). These studies have uniformly found that a wide range of video quality are acceptable to observers, but have not looked at what information people are able to extract for different levels of image quality.

Of those that have looked at whether image quality affects perception, there has been great variability in how image quality was varied. Studies by Ghinea and colleagues have generated fairly high quality, multimedia videos that include audio, video, and text, and have found that what is learned from the videos depends on the dynamism of the content but not on frame rate (varied from 5-25 fps) or whether the video was colour or B/W (Chen, Ghinea & Macredie, 2006; Gulliver, Serif & Ghinea, 2003). It is important to note that what was to be learned may have been conveyed by sound as well as by image, in this study.

The number of frames per second (10 vs 15 fps) affects the ability of lip-reading deaf adults to understand speech over a videophone that uses quite a bit of data compression that is greater in the periphery than in the centre of the picture (Woelders, Frowein, Nielsen, Questa, & Sandini, 1997). Speech comprehension is a highly sequential task, and so it is not surprising that removing the rate of information over time will impact comprehension. Comprehension of finger-spelling in the same study was not affected by the same manipulation of frames per second in the videophone. This makes sense when you consider that finger-spelling presents letters at approximately 4 letters per second, which are adequately sampled by a video progressing at 10 frames per second.

Yoo, Kim, Jun, Kim, Lim, and Kim (2004) tested the performance and perceived image quality of an ultrasound video transmitted over communication lines. Their images were 320 x 240 pixels, and their system was capable of transmitting 30 fps. What varied in the study was the communication line, which affected the degree of image compression, the amount of noise that appeared in compressed and then decompressed images (using a metric called peak signal to noise ratio), and the speed of image transmission, measured in Mbit/s. They found that radiologists who were to diagnose using the ultrasound images rated the quality significantly better for uncompressed images than for compressed images when the bit rate was less than 0.8 Mbit/s. When frame rate was added to the analysis, it interacted with bit rate in affecting image quality. They reported a threshold bit rate needed to achieve satisfactory image quality for each frame rate. The threshold bit rate was 4, 0.8, 0.6, and 0.6 Mbit/s for frame rates of 5, 10, 15, and 30 fps, respectively. Their analysis of the best possible distinction of targets from background noise achievable at the different bit rates and frame rates found that a satisfactory target/noise ratio was found as long as bit rate was at least 1, 2, 2, or 4 Mbit/s for frame rates of 5, 10, 15, and 30 fps, respectively. When they looked at both measures together, they suggested that a bit rate of more than 0.6 Mbit/s, at 30 fps should be the minimum to maintain diagnostic quality of the ultrasound images. This was beyond the capability of the ADSL communication lines tested, but was easily achieved by the VDSL and cable lines tested.

**Spatial resolution.** A number of studies have shown that if a person's actions are grossly simplified by recording their movements in the dark when they are wearing only small dim lights at various places on their bodies, observers of the resulting movies can identify both the person and the person's actions while the dots are moving, but not before the motion begins. Understanding of these *point-light displays* is fairly universal when the motion is of people or animals, with recognition of the action being better when it is a locomotion action rather than a social or instrumental action (Dittrich, 1993). Motion also improves recognition of the shapes of and sometimes the actions of inanimate objects (Anderson & Bradley, 1998). There are limits to the usability of such degraded displays, however. People who were asked to discriminate gestures that were part of a gestural language performed less well when looking at videos of point-light displays of the gestures rather than the full image of the videos (Herman, Morrel-Samuels, & Pack, 1990).

**Summary and implications for CCTV review.** Our understanding about how people identify and understand actions is fairly limited, but it involves both instinctual understanding of action organisation and learned understanding of intentional behaviour. There has been one study on detection of intentional action in CCTV footage (Troscianko, Holmes, Stillman, Mirmehdi, Wright & Wilson, 2004). The study considered CCTV surveillance rather than CCTV footage review. However, the study used real footage. The quality of the footage is not described, but example videos provided were B/W and relatively high quality. Both novices and experienced CCTV monitors were able to robustly detect the intention of people to act violently before the violence began, but were less able to detect the lead-up to intentional theft and criminal damage.

Actions can be detected in some degraded images, but identification of the actions is generally less accurate. Most research on this topic has aimed to understand the degree of noise and degradation in a multimedia system that still allows people to process for enjoyment and learning reasonably well. This research has found we tolerate video degradation up to a point. Further, in a study by Scott-Brown and Mann (2005), an unspecified group of participants were asked to look at CCTV footage and detect non-violent but suspicious behaviour. Detection was poorer overall than in the Troscianko et al (2004) study. In terms of the effect of image degradation, the Scott-Brown study compared high quality B/W footage to high quality colour footage and both to the same footage with a time-lag of 1 frame inserted between each frame. Detection of suspicious behaviour was better when using high quality B/W footage than when using high quality colour footage. The introduction of the time-lag, which would have severely disrupted perception of motion,

lowered detection in the B/W condition significantly, and seemed to be approaching such a drop for colour footage. It may be that performance was too low already in the colour condition to be able to pick up a further drop due to time-lag.

Depending on what actions CCTV reviewers are searching for, degradation may or may not have a negative impact. If the action is fast or a target appears briefly, a reduced frame rate will certainly diminish detection of the action. But if action is slower or a target appears for more than a few frames, the reduced frame rate may have little impact.

## **4 *Distractions and Interruptions***

### **4.1 Distractions**

When looking for targets, it seems likely that irrelevant information in the footage will draw attention occasionally. Certainly, in models of attentive behaviour, attention is usually characterised as being directed both by the goals people have as they begin to search and by the saliency characteristics of what people are looking at (e.g., Egeth & Yantis, 1997; Wolfe, 1994). Much applied research has been aimed at understanding what distracts people, both in order to help people fight distractions (e.g., designers of education materials) and to help designers capture attention when people are engaged in a task that is irrelevant to what the designers want people to do (e.g., designers of advertisements) (Hillstrom & Chai, 2006). Popular wisdom is that objects that appear abruptly or flash draw attention, movement captures attention, and distinctive or novel objects capture attention. In fact, recent models of how eye-movements reflect search behaviour as people look at displays have argued that salience in displays is the primary determinant of how people scan images (Itti, Koch & Niebur 1998). It turns out that in the laboratory, most of these things only capture attention strongly under limited situations (e.g., Hillstrom & Yantis, 1994), and so the conclusions most researchers have reached is that attention is more goal directed than our intuition leads us to believe. Attention and our eyes are captured only by those elements in the display that have the most salient stimulus characteristics relevant to our goals (Folk & Remington, 1998; Henderson, Brockmole, Castelano & Mack, 2007; Remington, Folk & McLean, 2001). In situations where we learn that the stimulus characteristics we seek will not be salient, we are capable of controlling our attention to ignore (or quickly reject) the most salient elements relevant to our goals in order to find elements with less salient features relevant to our goals. If this selectivity were not possible, search would fail more often than it does.

So why are we ever distracted by irrelevant but salient objects in the world? First, all the studies that showed people could successfully ignore irrelevant stimuli were relatively short in duration and were run in environments that purposely removed potentially distracting stimuli, other than the stimuli purposely included in the study. In those situations, people maintain goal-related behaviour well. However, it is difficult to maintain a goal for a long period of time. The skill required to do this is called executive control, and the section on Time on Task can be thought of as describing failures of maintaining goals. Secondly, our brain's initial assessment of which objects are task-relevant is quick and approximate, and so sometimes objects are wrongly prioritised as highly relevant to the correct goal.

Finally, and probably most importantly, it is a rare situation to have only a single goal influencing one's behaviour. We all have chronic and usually low-grade motivations for our behaviour that exist concurrently with the temporary goals we adopt in order to do tasks. Some distractions reflect an object in the world strongly activating a weakly chronic goal that is secondary to the task-relevant goal. An example is that smokers are more distracted by smoking-related stimuli than are non-smokers (e.g., Bradley, Mogg, Wright & Field, 2003).

Given the paucity of evidence that salient objects capture attention regardless of our goals, why do people believe so strongly that they do? Firstly, something not capturing attention does not reach our awareness, so our beliefs about what captures attention are biased by those situations in which our attention *is* captured, ignoring the more frequent times our attention *is not* captured. Secondly, salient things are more memorable than routine things (e.g., Rajaram, 1998), so episodes of attentional events come to mind more readily when something salient was involved than when nothing salient was involved. Finally, memorability has been shown to influence people's assessment of the probability of events (Tversky & Kahneman, 2002), and so this, too, biases us to believe salience is more important than it is.

## 4.2 Interruptions

When attention is moved away from the primary task because of distractions, either incidental or purposeful, the task has been interrupted. Resumption of the primary task takes measurable time (Trafton, Altmann, Brock & Mintz, 2003), although sometimes people are able to overcome the slow resumption by speeding up the primary task after it resumes (Zijlstra, Roe, Leonora & Krediet, 1999). Interruptions to simple tasks rarely result in performance errors on the primary task (Speier, Valacich & Vessey, 1999), even if the primary task involves multiple simple goals that are being coordinated and managed (Law, Logie, Pearson, Cantagallo, Moretti & Dimarco, 2004). Interruptions to complex tasks, on the other hand, can result in errors, especially if the nature of the interrupting task is similar to the primary task (Edwards & Gronlund, 1998; Gillie & Broadbent, 1989). Resuming an interrupted comprehension task is less effective when one can review only the high-level meaning of what was being comprehended than when one can review the most recent details presented before the interruption (Glanzer, Fischer & Dorfman, 1984). Finally, resuming the interrupted task is less effective when the interruption has to be processed immediately rather than when people can take a few seconds to transition from the primary task to the interrupting task (Altmann & Trafton, 2002).

The research about interruptions discussed in the preceding paragraph involved cognitive tasks that were not search tasks. There has been a flurry of research recently about interrupted visual search. Where the interruption is caused by inserting a blank screen for less than a second between two displays of a static search display, the time for resuming search is faster than the time to start search in the first place (Lleras, Rensink & Enns, 2005). The benefit seems to be caused by the establishment of the first, unconscious representation of the target long before it is consciously found (Lleras, Rensink & Enns, 2007). If the target is different after the interruption, responses are slowed, but if the nontargets are different, response times are unaffected (Lleras et al, 2007). Passive viewing of an unrelated scene during the interruption does not change the performance at resumption relative to passive viewing of a blank screen (Shen & Jiang, 2006). But if instead of passive viewing, the observer is interrupted to do a second search task, they lose track of their original search task (Shen & Jiang, 2006). The researchers' compelling interpretation of these findings is that what is critically disrupted by a second search task, but not by passive viewing, is memory for spatial layout of the first display. It should be noted when considering this research that the displays were clear and used quite distinct and artificial objects. It is not at all clear what would happen if the displays were degraded or if they showed naturalistic scenes rather than arrays of geometric objects.

**Summary and implications for CCTV review.** Although salient objects do not inherently draw attention, they are more likely to be attended when they are similar to task-related goal or are strongly relevant to a secondary motivation of the reviewer. Interruptions caused by distractions or by purposeful requests that the reviewer do something else (momentarily or for a long time) may not cause problems if the reviewer is doing a simple search, but will be more detrimental if the reviewer is doing a more complex review of the footage. To avoid distractions while reviewing CCTV footage, the environment should be clear of unnecessary distractions and the reviewer should not have his or her mind on other tasks or goals.



## **5 Time on task**

### **5.1 Maximizing comfort over long periods of time spent reviewing footage**

Part of the problem of long work viewing footage is that it can create discomfort and, if repeated a lot, injury to the person viewing. Advice about workstation setup includes the following points (Health and Safety Executive, 2008; US Department of Labor, 2008):

- Room lighting should be relatively bright so that there is not sharp contrast between the monitor and its background. Viewing such contrast for a long time can tire the eyes. Less bright lighting can be used for LCD monitors than for CRT monitors.
- Room lighting should be diffuse to help avoid spots of glare.
- Blinds should be available for windows.
- Monitors should be cleaned of dust and grime, as it can increase glare.
- If there is a window or other side light source, the monitor should be placed to neither reflect on the monitor nor sit between the viewer and the light source. Generally, monitors should be placed with their face perpendicular to side windows.
- The person should not sit under air conditioning or heating vents unless the vents are designed to redirect the air flow away from the person. Otherwise, eyes can dry. It is better if air flow is between 7.5-15 centimetres per second. It is best if humidity is between 30% and 70%.
- The US recommended indoor air temperature is between 20° and 23.5° C during heating season and between 23° and 26° C during the cooling season. The UK recommended that indoor air temperature is over a minimum of 16° C, and this difference no doubt reflects differences in what people in these two countries are accustomed to. What is important to keep in mind is that review of CCTV footage requires people to be stationary for long periods of time, and so the room should not be too cold.
- There should be a way to bring down the temperature if the room gets hot. Fans or windows suffice.
- The centre of the monitor should be approximately 15° below horizontal eye level, to avoid neck and back strain. Depending on the size and distance of the monitor, this is usually achieved by having the top of the monitor be at about forehead level.
- Chairs should be adjustable, and should provide back support and allow the seated person to keep their feet on the floor and sit up straight.
- The seat should be low enough that it does not cut into the back of the person's legs.

Although these guidelines should be followed when setting up any computer workstation, the concentration required to detect targets in degraded images will magnify any eyestrain that a poor setup produces. Further, discomfort can be distracting when people are having difficulty concentrating.

## **5.2 Maximizing performance over long periods of time spent reviewing footage**

Aside from the health and safety implications of long hours spent reviewing footage, there are performance implications as well. In particular, time on task affects vigilance, or readiness to react. Most tasks used to study vigilance are tasks in which people are monitoring an information source for targets, and must react in some way to the targets, when they appear. The review of CCTV footage is such a task.

There is a well known relationship between the time spent doing a task that requires vigilance and the quality of that vigilance. Overall, vigilance usually degrades over time. For instance, well trained and highly motivated radar operators were 10% less accurate after 30 minutes of watching radar than at the start of the watchperiod (Mackworth, 1948). This has led most employers who require their employees to maintain a high level of vigilance to institute mandatory breaks from the monitoring task. The resulting “watchperiods” (length of time vigilance is to be maintained) range from 20 minutes to a few hours (Angus, 1984).

Vigilance is affected by the speed at which information is arriving, by whether there are multiple sources of information or a single source, and by whether targets are detected by comparing them to simultaneously presented non-targets (“simultaneous task”) or by judging them in isolation based on a maintained mental target “template” (“successive task”) (Davies & Parasuraman, 1982; Parasuraman & Davies, 1977). Vigilance can be maintained longer when targets are distinctive against the background than when they are difficult to see (See, Howe, Warm & Dember, 1995; Warm, 1984). It is more difficult to monitor multiple information sources than a single one. It is more difficult to maintain vigilance when the task is successive rather than simultaneous, when information is arriving faster rather than slower, and when information to be assessed is expected to arrive at unpredictable times rather than at a fixed rate. The rate and regularity of information arrival is more critical when the task is successive than when it is simultaneous (Warm & Dember, 1998). Note that all but the speed factor are at their non-optimal when reviewing CCTV footage.

Vigilance usually degrades over time, but what happens cannot be described simply as poorer performance. The kinds of errors that could be made are mistaking targets for non-targets or mistaking non-targets for targets. What often happens as time on task increases is that people miss targets more often but make fewer misidentifications of non-targets as targets (Warm & Dember, 1998).

Originally, vigilance was thought to wane because monitoring is boring, and so arousal is low. In simple monitoring tasks, this may be the case (Pattyn, Neyt, Henderickx & Soetens, 2008). However in more demanding tasks, the drop in vigilance over time is more likely due to the very high cognitive load that such tasks place on people. Constant observation and decision-making is very demanding, and so people can only do it well for a limited period of time (Warm, Dember & Hancock, 1996). Theorists have speculated that the reason successive tasks show more of a vigilance decrement than simultaneous tasks is that successive tasks require maintaining an accessible mental representation of the target, which is demanding, and recent research suggests that it is this, rather than low arousal, that causes poorer vigilance with longer time on task (Caggiano & Parasuraman, 2004).

**Controllable factors affecting vigilance.** Other factors, too, affect vigilance, and indeed how long people can maintain their vigilance. For instance, it is easier to maintain vigilance when one is not tired (Anderson & Hone, 2006; Krueger, 1989). Caffeine can help people counteract the effects of sleep deprivation on vigilance (Lieberman, 2003). Short naps can help, but it is important to recognise that it can take up to 20 minutes to reach full alertness after a nap (Krueger, 1989; Craig, 1984).

Kjellberg (1990) reviewed the literature on how noise affects performance. He points out that research suggests that constant noise tends to increase fast errors in simple vigilance tasks, and also leads to a faster waning of vigilance. Constant noise increases the monotony of the environment and leads people to respond more slowly to rare stimuli and react more slowly to changing stimulus probabilities. More variable noise can be annoying, but it interferes most when it has information content (e.g., overheard conversations). Overheard conversations are most disruptive when the task being done requires use of verbal short-term memory. Noise can help to mask this kind of annoying, potentially distracting noise in the periphery. Finally, variable or informative noise (e.g., music or talk over radios) can sometimes help maintain arousal when people are getting drowsy from a monotonous task (Craig, 1984). Turning on or adjusting the radio was highlighted as the common technique drivers use to increase arousal when driving while sleepy, but there is no scientific evidence that it helps (Nguyen, Jauregui & Dinges, 1998). Nguyen et al did describe one study that

found a small benefit of the radio for extroverts, as compared to introverts, but concluded that it is more common for research to find a decrement in performance from the radio rather than a benefit.

Although the results of individual studies varied substantially, a meta-analysis of studies found that very hot and very cold environments can also negatively impact vigilance (Pilcher, Nadler & Busch, 2002). Small changes in skin temperature can impair vigilance through distraction (cooling: Cheung, Westwood, & Knoz, 2007; warming: Raymann & Van Someren, 2007).

It is easier to maintain vigilance when one is comfortable. Chronic pain impairs vigilance (Von Bueren, Radanov & Jäncke, 2005), and even trivially mild discomfort can reduce vigilance, presumably through distraction (Bell, Cardello & Schutz, 2003). On the other hand, introducing a moderate amount of discomfort has been shown to increase vigilance over the short term (Craig, 1984), as it can increase arousal when arousal is dropping.

**Summary and implications for CCTV review.** As time on task increases, people miss targets more often, particularly when information is arriving from more than one source, when what one is searching for has to be maintained in the searcher's memory, when information arrives fast and/or unpredictably, and when the likelihood of finding a target is low. CCTV footage review has many of these characteristics. In addition, there are situations in which vigilance does not wane over time. If a task is so practiced that it becomes automatic behaviour, vigilance tends not to decrease with time on task (Fisk & Schneider, 1981). However, CCTV footage review is not likely to become an automatic task, because the nature of the target being sought changes for different investigations.

Further, there are some situations in which vigilance improves with time-on-task. This is likely to occur only when the task requires (a) interpretive judgments of stimuli (e.g., "could be a weapon") rather than perceptual judgments of stimuli (e.g., "is blue") (b) made by comparing the stimulus to its background, and (c) the rate of information arrival is slow (Warm & Dember, 1998). In this case, observers are likely to be learning how to best make their judgments as they get more perceptual experience. Within a single investigation, information arrives relatively slowly and interpretative judgments are common, so some learning is likely to be evident before any vigilance decrements are apparent.

So will CCTV footage review show vigilance decrements when people engage in it for long stretches of time? To the best of our knowledge, there is only one study that has looked for

vigilance decrements. When people reviewed CCTV footage for 4 hours, with one 10 minute break in the middle, performance improved over time rather than degraded (Tickner & Poulton, 1973). However, it is important to note that (a) evidence suggested that people were still learning what their targets looked like, and (b) the test for vigilance decrement was run using people inexperienced in reviewing CCTV footage.

There are some people who argue that vigilance decrements happen far more in laboratory research than in the workplace (e.g., Nachreiner, 1977). It is true that field studies of vigilance have found weaker effects than laboratory studies. There are a number of reasons for this. First, participants in laboratory studies are often less motivated to do their task than people doing a job in the workplace. Second, workplace tasks are often more complex than the tasks used to study vigilance in laboratories. And finally, in the workplace, people often either take brief breaks periodically or vary what tasks they do over time, so that they are in essence managing their alertness themselves. But even with these differences, there have been a number of demonstrations of vigilance decrements over time in the workplace, for instance in monitoring of radar (Pigeau, Angus, O'Neill & Mack, 1995), in air traffic control centres, in luggage search in airports, in monitoring patients' vital signs during medical operations, in driving cars, lorries, and trains, and in teamwork in simulated battle (Harville, Lopez, Elliott & Barnes, 2005). Thus, while it is yet to be determined the degree to which time on task affects CCTV footage review, there is good reason to consider that it might.

There are times when CCTV footage review is better described as a task requiring integration of a flow of information rather than detection of a discrete target. Such would be the case when trying to develop a timeline of events leading up to a crime. The direct relevance of the vigilance literature is a bit more questionable in these cases. Nevertheless, there have been some reports of vigilance decrements in more complex tasks than simple target detection (Weinger & Smith, 1997), and so it is prudent to consider whether vigilance could be an issue in even more complex review of CCTV footage.

**Improving vigilance.** The ability to control the pace of work and the ability to take short breaks improves resilience to lapses of vigilance (Krueger, 1989). Reviewing CCTV footage falls somewhere between the extremes of machine-paced work versus worker-controlled tasks. The reviewer is free to stop the footage at any time, but the pace of the footage is not under control except during those breaks. And stopping the footage can lead to interruption effects, discussed in the Distractions and Interruptions section of this report.

Providing CCTV reviewers a work environment that has a comfortable temperature, or at least the ability to open a window or run a fan, is likely to help the reviewer manage their own vigilance to some degree. Likewise, the ability to use a radio when vigilance drops or when conversations in other rooms are distracting may help.

**Monitoring vigilance.** Is there a way to detect that someone's vigilance is dropping to unacceptable levels during the task? The most common techniques used in industry are to insert artificial images of targets into the images being inspected, or real objects that include a known proportion of targets. Not only does this allow periodic checks of vigilance, it also offers the opportunity to give periodic feedback about performance. However, this is unlikely to be a workable intervention in the context of reviewing CCTV footage, not least because altering the footage might compromise the evidentiary value of the footage. Furthermore, the equipment being used to review footage is so variable currently, it is difficult to envision a robust way to superimpose targets onto footage without altering the footage.

If vigilance needs to be monitored, it might be more effective to use some physiological monitors for arousal levels. For instance, a recent article reports a monitor worn on the wrist that assesses arousal and vigilance (Lieberman, Kramer, Montain & Niro, 2007). Such a device could, in theory not only monitor alertness but also sound an alarm if vigilance is falling below acceptable levels. Other approaches have been to monitor EEG functioning (e.g., Wilson & Russell, 2003) or eye movements (e.g., Marshall, 2007).

**Reducing time on task by reviewing multiple tapes at once.** It is no doubt tempting at times to try to review more than one tape at a time. The only studies found on monitoring more than one tape at a time were focused on monitoring CCTV rather than reviewing it. Nevertheless, some of the outcomes are relevant to the review of footage.

In one such study, people were asked to monitor up to 16 television monitors showing scenes inside and around a prison. This research found that when there was almost no movement in the displays, people missed only 3% of moving targets when monitoring 16 screens (Tickner, Poulton, Copeman, & Simmonds, 1972), but when there was substantial movement in the displays, performance dropped substantially (Tickner & Poulton, 1973). The latter study showed scenes inside and outside prisons, with the targets being certain events that people did (e.g., throwing an object, parking in a certain place), and participants were either university staff and students, experienced prison guards or inexperienced prison guards. Performance ranged from 59%-63% detection of targets for the three groups, with detection inside the prison, where experience was relevant, ranging from 87% for

experienced guards to just over 70% for the others. Other conditions, which used only university participants, showed that having to answer a telephone periodically, but never less than a minute before an event occurred, dropped detection rates a small but reliable amount. Detection improved when fewer television screens were watched (76% for nine screens, 84% for four screens, 100% for one screen). Performance was better when screens all showed the same kind of scene (e.g., all inside the prison) than when half showed one kind of scene and half another (e.g., half inside the prison and half outside).

In another study, Tickner and Poulton (1968) had people monitor a bank of television monitors to detect cars needing assistance on the side of the motorway. This was an example of searching for targets that may be difficult to see but that are not as brief as in the prison observation study. Observers looked for parked cars on the shoulder of a motorway that could be seen clearly or barely at all (due to distance) on video showing the output of one or three CCTV cameras. When there were three cameras, in one condition the image changed every 10 seconds, in a second condition it changed every 10 seconds unless the observer pressed a button to stay on the same camera for an extra 10 seconds, and in a third condition the monitored changed only when the observer chose to change it. A group of police officers were faster than civilians at detecting near targets and more accurate at detecting far targets, when both were monitoring three cameras simultaneously. All other conditions used only civilian observers. Monitoring three cameras simultaneously produced comparable detection rates to monitoring only one camera, but responses were slower. When monitoring three cameras, responses were slower when only one camera's image was presented at a time than when all three were shown simultaneously, particularly for nearer targets. Responses for distant targets were more accurate when the observer could extend viewing time than when the image cycled between cameras automatically.

In a third study, Tickner and Poulton (1975) had observers looking at a single television monitor to detect target people walking down a street, while simultaneously monitoring for criminal actions (theft and transferring goods) of anybody in the video. The target people were indicated by posting photos of them near the television screen. The majority of conditions used a four hour video of a single street, but in two one hour conditions, a colour video was compared to the same video in black and white. The video showed a street extending straight ahead into the distance, and so the distance of actions varied and the distance at which people approaching could be identified was one of the measures taken. For the 4 hour video, police officers were slightly more accurate at detecting theft. They were approximately equal at detecting the target people, and showed slightly fewer false alarms on misidentifying people and actions. All participants were able to monitor for one or three

target people equally well. When viewing the one hour B/W film, detection of target people was best when monitoring for three people, slightly worse when monitoring for seven people, and considerably worse when monitoring for twelve people. Detection was better when monitoring the colour film for three people than when monitoring the B/W film for three people. Far actions were detected less often than near items. When a target person committed a crime, the crime was detected more often than the person was. When the same target person committed a crime both in the near distance and the far distance, they were detected as target people more often the closer they were.

A follow-up study (Simmonds, Poulton & Tickner, 1975) showed that when observers searched a one hour B/W video for five people, their performance was much better if the photos of the target people showed them wearing the clothes worn in the video rather than other clothes. In addition, detection required the target people to be about twice as close as when they were seen in daylight conditions.



## **6 Mindset**

The ability to use CCTV footage effectively depends on maintaining an appropriately open and pragmatic mindset when approaching the review of footage. We have already pointed out that expectations about what will be seen strongly affect attention. It also affects decision making. Approaching tasks with a knowledgeable and goal-oriented mindset is a skill that generally distinguishes expert behaviour from novice behaviour. On the other hand, too strong an expectation of what one is looking for can also lead to biased decision making.

### **6.1 Maintaining the ability to detect unexpected things**

In terms of how our attentional system aids our behaviour, it is just as important that our attention can be captured as it is that we can maintain focus. Our attention is always engaged in some task, so without attention capture, nobody could ever get us to pay attention to the appearance of danger. In terms of reviewing CCTV footage, if a reviewer is engaged in searching for something in particular and something else appears that ought to open the possibility that ought to be considered important, the reviewer would not notice it if too engaged on the original search task. It is vital that we notice things that occur and that we are able to disengage from a current task.

The more difficult our primary task is, either because of perceptual complexity or because of the need to maintain information over time while we are engaged in it, the less likely it is that we can be interrupted by the appearance of something new and potentially important (Lavie, 1995; Lavie 2005). Further, alarms meant to draw our attention work better if they appeal to multiple senses rather than only one (Santangelo & Spence, in press). And just as it is true that perceptually salient objects only reliably capture our attention if related to a goal or motivation we currently have, it is more likely that we will notice changes if they are related to our goals and setting (Most, Scholl, Clifford & Simons, 2005).

The laboratory research on this in the perceptual domain focuses on two related phenomena: inattention blindness and change blindness. Inattention blindness is the failure of people to notice an unexpected, task-irrelevant stimulus when focused on a goal, even when the stimulus is within the spatial bounds of what is attended. Detection of unexpected stimuli has been estimated at anywhere from 10-50% for simple stimuli that would be easy to detect if one were looking for them, with the rates increasing when the stimuli have meaning or bear a perceptual relationship to what is being attended (Mack & Rock, 1998; Simons & Chabris, 1999). Once people are alerted to the possibility that there might be such an

appearance, detection rates rise substantially. This phenomenon has been investigated using not only simple geometric stimuli (e.g., Rock & Mack, 1998), but also photographs of scenes (e.g., Hollingsworth & Henderson, 2000), and dynamic movies or real-world scenes (e.g., Levin & Simons, 1997; Simons & Chabris, 1999; Simons & Levin, 1998).

Change blindness is the failure of people to notice perceptual changes even when they are looking for them. It is most commonly measured as the time it takes for people to detect the change when people are shown alternating original and changed images repeatedly (Rensink, O'Regan, & Clark, 1997), with some interruption between the images (e.g., a blank screen) so that detection relies on comparing immediate perceptions to working memory. Again, once the change is detected, it usually seems laughably noticeable, but the failure or slowness to detect such changes is a robust phenomenon that demonstrates how much less we see than we think we see. What may be even more surprising than the failure to detect changes is how unaware we are of this failure. Even after being made aware of the condition, people typically underestimate the degree to which they, individually, will be subject to it (Levin, Momen, Drivdahl & Simons, 2000).

For the purposes of this report, it is fair to discuss these two phenomena as one, and so the label “change blindness” will refer to both henceforth. Investigations into how change blindness works has found four critical findings: (1) Observers who expect that they may be shown a change are more likely to detect it than observers for whom the change is completely unexpected (Mack & Rock, 1998). (2) When deliberately searching for changes, changes to an object in a scene are only detected if they occur while the object is attended (Scholl, 2000). (3) Changes to elements of a scene important to the meaning of the scene are detected more than elements of a scene that are incidental to the scene's meaning (Hollingsworth & Henderson, 2003). (4) When there is action in what is being watched, changes at transitions between actions (e.g., between reaching for a suitcase and picking it up) are noticed more than changes in the middle of actions (Baird, Baldwin & Malle, 1999, as reported by Saylor & Baldwin, 2004).

Research by Simons and Levin (1998) is particularly relevant. In one study, they set up an experiment run on unsuspecting people on a university campus. The experimenter walked up to a student pedestrian and asked for directions. Two confederates of the experimenter then walked between the experimenter and the person being studied carrying a large door, and while the board passed, the first experimenter was replaced by a second experimenter who had been behind the board. The two experimenters were both men and were not extremely different in appearance, but they had quite distinctive voices, were 5 cm different

in height and wore different clothes. Fewer than half of the pedestrians noticed the change in who they were giving directions to, despite the fact that most made eye-contact with both the first and second experimenter during the course of the interaction. Detection was more common when the experimenters were in the same cohort (i.e., were approximately the same age and were dressed similarly to the person giving directions) than when the experimenters were younger than the person giving directions or were dressed so as to be clearly non-students. Although it is tempting to dismiss this research as less relevant because the people who were asked to detect the change were not focused on identifying the experimenters, we would refer back to the difficulty people have in detecting changes that they are trying to detect. Levin and Simon's (1997) work using film clips deliberately had continuity errors between shots (e.g. a scarf worn by an actor in one shot was not being worn in the next). In the condition in which the observers did not know that there might be changes, they seldom noticed any. But even in the condition in which people were told to report all the continuity errors they could find, people detected only approximately 22% of changes.

Would the same problem be found if people were looking at a single object to see if it changed? In the Levin and Simon (1997) study, everybody could detect a change in actor if told to pick out movies in which actors changed. But there is a substantial body of research asking people to judge whether two images show the same object (in direct orientations) or different objects. In such studies, familiar objects are judged more slowly when the objects are shown in different orientations (e.g., Lawson & Humphreys, 1996), and unfamiliar objects are judged less accurately when the objects are shown in different orientations (Tarr, Williams, Hayward & Gauthier, 1998). It is reasonable to assume that image degradation would lead people to act towards familiar objects more as if they were unfamiliar objects.

**Implications for CCTV footage review.** The relevance of change blindness to surveillance and monitoring situations has been recognised for CCTV surveillance (Scott-Brown & Cronin, 2007), for military traffic control rooms (DiVita et al, 2004)), for cockpit displays (Haines, 1991) and for other forms of military control equipment that need to be monitored (Durlach, 2004).

The change blindness literature is relevant to *review* of CCTV footage in two main ways. First, it strongly implicates the role of expectancy in the search process. Observers who are aware of their expectations as they search and can flexibly work with alternative expectations are more likely to detect out of the ordinary targets than those with more rigid cognitive processing. Secondly, it provides a caution to check one's assumptions when

assuming one sees continuity of objects, people or events from one camera's footage to another. The expectation built up from one bit of footage could well bias the reviewer of footage in how they interpret what they see in the next bit of footage.

## **6.2 Mindset in decision making**

The success of an investigation depends to a large extent on the ability of the investigators to evaluate information and evidence accurately (Ask, Rebelius & Granhag, 2008).

Identifying, extracting and drawing accurate inferences from CCTV evidence requires the same accuracy and relies on the ability of investigator to remain objective and open to alternative interpretations (or possibilities) when reviewing and evaluating such evidence.

Ideally, the evaluation of evidence should not be affected by external or contextual factors, such as time pressure, preconceptions, emotions, beliefs about likely suspects or sequence of events. Indeed, it is generally assumed that we make 'hard-nosed', objective decisions and judgments as a matter of course and that, in particular, experienced individuals will not be influenced by the vagaries of contextual factors. However, the results of research challenge this assumption and suggest that the evaluation of forensic evidence can be sensitive to external influences (Ask, Rebelius & Granhag, 2008; Ask & Granhag, 2007; Dror, Charlton & Péron, 2006; Dror & Rosenthal, 2008).

At the outset of an investigation, police will be guided in their search and evaluation of evidence by preliminary or working hypotheses concerning the crime. For instance, how the crime was committed, who was involved and why it occurred. These working hypotheses may not be based on available, objective evidence – quite often, evidence of that type may not be available. Rather, these hypotheses may be based on expectations or script-based causal explanations (Ask & Granhag, 2005). In other words, investigation is hypothesis-driven as investigators try to piece together any available evidence to formulate the most plausible account of the crime. Constructing theories which provide a causal structure for information or evidence is frequently a spontaneous cognitive response to ambiguous problems or situations (Kahneman, Slovic & Tversky, 1982; Nisbett & Ross, 1980).

In any investigation involving CCTV evidence, it is quite likely to be the case that the officer trawling through CCTV footage will also have access to other details about the case and may be working with quite specific hypotheses concerning who and what to look for among many hours of CCTV recordings. Ideally, these hypotheses will be based on the triangulation of evidence from other sources, such as victims, witnesses, informers and so

on. However, human cognitive processes are such that preconceptions, expectations, pre-existing schemas for particular crime types and other biasing tendencies may (unhelpfully) influence the evidence-evaluation process. These tendencies are fundamental to human information processing and have been documented throughout the psychological literature.

The following section will examine several of these tendencies as they relate to the evaluation of forensic evidence.

**Confirmation Bias.** Confirmation bias refers “to unwitting selectivity in the acquisition and use of evidence” (Nickerson, 1998, p.175). In other words, the tendency to favour information or evidence which confirms an initial or existing belief while avoiding or rejecting disconfirming evidence (Koriat, Lichtenstein, & Fischhoff, 1980).

Research by Darly and Gross (1983) provides a good illustration of the impact of prior expectations and preconceptions on ability to objectively evaluate actual evidence. In their study, two groups of people viewed a videotape of a child taking an academic test (the same child was viewed by both groups). One group was led to believe that the child came from a high socioeconomic background while it was suggested to the other group that the child’s socioeconomic background was low. Both groups were asked to rate the academic ability of the child based on what they had seen of their performance in the video alone. Participants in the former group (high SES) rated the child’s abilities higher than those who were led to believe that the child came from a low socioeconomic background. Darley and Gross (1983) argued that participants formed an advance hypothesis about the child’s academic abilities on the basis of socioeconomic background and then unwittingly sought out evidence in the video recording that was consistent with this hypothesis.

Research consistently demonstrates that we prefer information biased towards our pre-existing beliefs or expectations (Hope, Memon, & McGeorge, 2004; Jonas, Shulz-Hardt, Frey, & Thelen, 2001) and attitudes, stereotypes and preferences (Lundgren & Prislin, 1998). Confirmation bias has been demonstrated relating to stereotypes about ethnicity (Duncan, 1976), clinical outcome (Swann, Giuliano, & Wegner, 1982); education (Foster, Schmidt & Sabatino, 1976) and gender (Oakhill, Garnham & Reynolds, 2005)

The confirmation bias has also been robustly demonstrated across decision making in diverse domains from formal problem solving (Wason, 1968) to social interactions (Snyder & Swann, 1978) and across real life domains including public policy rationalisation (Tuchman, 1984), medical decision making (Elstein, Schulman & Sprafka, 1978) and judicial reasoning

(Hope, Memon & McGeorge, 2004; Kalven & Zeisel, 1966; Pennington & Hastie, 1986, 1988, 1993).

More recently, research has examined how confirmation bias impacts on the evaluation of forensic evidence by experts. Early studies showed that even the interpretation of visual evidence could be biased by expectations. For example, Bruner and Potter (1964) showed participants a set of blurred images which were gradually brought into focus. They found that exposure to extremely out of focus images made it more difficult for participants to identify the image as it was brought into focus (i.e. early interpretations of the image inhibited subsequent correct recognition). This phenomenon has been replicated on numerous occasions with the same results suggesting that the initial hypotheses that people form to explain or understand ambiguous event may make it difficult for them to interpret subsequent detailed information.

In a series of studies examining the accuracy of fingerprinting experts Dror and his colleagues (see Dror & Charlton, 2006; Dror, Charlton & Person, 2006; Dror, Peron, Hind and Charlton, 2005) found that fingerprint matching decisions, including those made by expert forensic examiners were also biased by extraneous contextual information. Specifically, that visual information (fingerprints) was interpreted in a manner consistent with initial expectations. For example, in Dror et al. (2005) the difficulty of the matching task was varied and some participants were also given additional information about the crime, such as where the fingerprint was obtained. Some participants also saw emotional photographs that related to the scene of crime. Finally, some participants were subliminally primed with the words “guilty” and “same” during the matching task. Results indicated that both emotions (as aroused by background story and photographs) and subliminal messages did influence decision making in certain circumstances. Specifically, when the matching task was easy (i.e. the fingerprints were a clear, uncomplicated match) the extraneous contextual factors did not affect the accuracy of the decision-making. However, when the task was difficult and the fingerprints were not a clear unambiguous match, errors consistent with the contextual information were observed. When the fingerprints presented were ambiguous, participants in the control condition found a match for 47% of the trials whereas participants in the high emotion plus subliminal message condition found a match for 66% of trials (58% in high emotion only condition). Dror et al. (2005) concluded that top-down influences (i.e. contextual information) biased decision making when the task was ambiguous but did not over-ride bottom-up processing (i.e. the objective analysis of fingerprint attributes) when the task was clear-cut.

Dror et al. (2006) replicated these findings in a similar study using fingerprint experts. In this study, the experimenters selected fingerprints that had previously been evaluated by the experts in the normal course of their work. These fingerprints were then submitted for a second analysis by the same experts (who were not aware which fingerprints they would be tested on or when). Participants were asked to examine the target fingerprint alongside an exemplar print (a print obtained from a suspect) and were then provided with inaccurate contextual information concerning the print. The misleading information was designed to generate an expectation that the print was a non-match – participants were told that the print had been erroneously matched by the FBI as the Madrid bomber. Results indicated that fingerprint experts, with on average 17 years experience, were just as susceptible to extraneous contextual information as non-experts. When presented with a different context, four out of five experts made different identification decisions to those they had made previously. In fact, three of those four experts decided that the fingerprint was now a definite non-match despite having identified those same prints as a definite match previously in the absence of contextual information.

Clearly this research has important applications, given that in reality, fingerprints – like CCTV images – are unlikely to be perfectly clear and obtained under optimal conditions.

Several other studies have recently investigated the impact of the confirmation bias and associated effects on evidence evaluation in forensic settings. For example, Ask and Granhag (2005) presented both experienced police investigators and a student sample with case materials relating to the preliminary investigation of a homicide. Background information was also provided which suggested that either the suspect had a motive or that an alternative unknown offender committed the crime. Neither hypothesis had any basis in available evidence. Results indicated that the student sample demonstrated a clear confirmatory bias – participants who were made aware of a potential alternative perpetrator were less likely to view the main suspect as guilty. In this instance, police investigators did not appear to interpret the evidence in line with the background information concerning an alternative perpetrator. Instead, this group rated it likely the main suspect was guilty irrespective of background information. Ask and Granhag (2005) suggest this may be due to another commonly held preconception or ‘guilt’ bias which has been documented elsewhere (Baldwin, 1993; Leo, 1996).

Research has also shown that the sequential evaluation of different pieces of evidence can be distorted in favour of an initial hypothesis. Hope, Memon and McGeorge (2004) tracked the course of evidence evaluation in mock juror decision making and found that the effect of

biased evidence evaluation was cumulative. Specifically, when an initial evaluation is biased (on the grounds of preconceptions or expectations) each subsequent evaluation biased in the favour of the previous (biased) evaluation. For jurors exposed to negative information about the defendant, this distortion process was exacerbated and the prosecution was more strongly favoured as the leader. In addition, evidence supporting the prosecution's case was more favourably evaluated or, alternatively, the evaluation of prodefense testimony was distorted in favour of the prosecution.

Other research has also demonstrated that the confirmation bias is exacerbated by sequential information processing (Jonas, Schulz-Hardt, Frey, & Thelen, 2001) and that people use different cognitive processes when faced with sequential versus simultaneous information (Hogarth & Einhorn, 1992). Under sequential presentation, new items are immediately compared with prior or pre-existing beliefs (Edwards & Smith, 1996) and assessed relative to this prior belief. Jonas et al. (2001) have argued that sequential presentation involves a repeated consideration of a prior belief or evaluation and a concomitant increase in confidence in the veracity and reliability of this prior evaluation. This repeated, but biased, evaluations leads in turn to increased commitment to the belief or evaluation (see also Koehler, 1991; Schulz-Hardt, Frey, Lüthgens & Moscovici, 2000; Tesser, Martin, & Mendolia, 1995).

Decision-makers who have been biased by their expectations or preconceptions are typically unaware that their decisions have been distorted and tend to retain an "illusion of objectivity" despite their selection attention to particular information (Pyszczynski & Greenberg, 1987). In other words, decision makers tend to report that their decision making has been unbiased and objective.

**Asymmetrical Scepticism.** A second related tendency which can influence evidence evaluation is asymmetrical scepticism. Rather like confirmation bias, asymmetrical scepticism is a naturally occurring tendency for people to scrutinise information which threatens previously held beliefs or preconceptions more rigorously than information which is positive with respect to an existing belief or value (Ditto, Scepansky, Munro, Apanovitch, & Lockhart, 1998; Ditto, Munro, Apanovitch, Scepansky & Lockhart, 2003; Lord, Ross & Lepper, 1979). Ask and Granhag (2007) examined the occurrence of asymmetrical scepticism in criminal investigations and found that experience interviewers judged the reliability of witness statements different depending on whether the statements confirmed – or disconfirmed – prior hypotheses held by the investigators concerning the case. Specifically, the results indicated that although the witness statements were produced under



the same circumstances by witnesses with the same characteristics “investigators subjected the disconfirming (vs. confirming) statement to stricter scrutiny and hence found stronger grounds for questioning its reliability” (Ask & Granhag, 2008).

In their most recent examination of bias in the evaluation of criminal evidence, Ask, Rebelius, and Granhag (2008) considered the extent to which different types of evidence might be susceptible to extraneous biases. Specifically, they considered the extent to which the perceived ‘elasticity’ of the evidence varied with objective nature of the evidence where ‘elasticity’ refers to the degree of ambiguity associated with the piece of evidence. In other words, if a piece of evidence is open to multiple interpretations, decision makers tend to give weight to the interpretation most consistent with their initial hypotheses or preferences. In the Ask et al. (2008) study, police trainees were given case materials (homicide) which contained information suggesting that a particular suspect was guilty. The purpose of this information, as in previous studies, was to set up a prior belief or hypothesis concerning that particular suspect. After participants had indicated their views on the case, new evidence was then introduced which was either consistent or inconsistent with the guilt-related suggestion. Furthermore, this evidence varied in terms of its perceived elasticity. In one condition, participants were provided with DNA evidence which is typically associated with a low degree of elasticity given the limited possibilities for subjective lay interpretations. In a second condition, participants were provided with a visual image (pictures taken from a CCTV security camera) which was deemed to have a moderate degree of elasticity (i.e. a moderate possibility of subjective interpretation). In the final condition, participants received details of witness evidence which was ascribed a high degree of elasticity on that grounds that witness evidence can be open to a number of biases and different interpretations based on both witness and situational factors (Wells & Olson, 2003). Analyses revealed asymmetric scepticism in the evaluation of evidence - participants rated the disconfirming (as opposed to confirming) evidence as less reliable and generated more arguments supporting this point of view. Furthermore, this scepticism was exacerbated for highly ‘elastic’ witness evidence.

**Other potential sources of biases.** A variety of stressors can affect decision making (Dror, Busemeyer, & Basola, 1999) and also increase the chances that biased information will impact negatively on judgements and decision making (Kruglanski & Freund, 1983). For instance, time pressure has been consistently shown to impair decision making and lead to increased bias and selectivity information processing (Edland & Svenson, 1993). Time pressure has also been shown to increase reliance on stereotypes and heuristics allowing less time for careful and considered processing of evidence (Bodenhausen, 1990; Dijker &

Koomen, 1996). Ultimately, time pressure can severely restrict the ability to generate and consider alternative hypotheses, thereby limiting the flexibility of the decision maker to consider other outcomes or possibilities.

**Summary and implications for CCTV review.** The implications of these findings for the review and analyses of CCTV footage are self-evident. Clearly, those engaged in the review activities should be very aware of the ways in which their decision making can, unwittingly, be distorted. Participants in all the studies reviewed were unaware that what they perceived as their 'objective' decision making had been manipulated and biased as a consequence of their pre-existing expectations, beliefs or hypotheses.

The uncertainty experienced by officers reviewing CCTV footage – uncertainty which can lead to error – is a common feature of decision making in the real world. There are a number of strategies available to decision makers to handle this type of uncertainty. The most obvious strategy is to reduce or remove the uncertainty. This might be achieved by collecting additional information or delaying further decisions until additional information comes to light (Dawes, 1988, Hirst & Schweitzer, 1990; Janis & Mann, 1977). However, this tactic may not always be possible. In an investigation, there may be no further information available until evidence obtained from CCTV produces a 'break' or identifies some important component of the incident. Thus, although decision making can be biased by expectations and pre-conceptions in the manner reviewed above, it can be possible to make use of assumption-based reasoning under informed conditions (Cohen, 1989; Lipshitz & Ben Shaul, 1997). This section will briefly examine decision making under uncertain conditions and how optimum decisions can be promoted in those circumstances.

Early decision-making theories typically proposed a deliberate rational decision-making process. Such rational theories are limited in their ability to explain decision-making in real-world contexts given the factors typically associated with real world decisions: time pressures, uncertainty, vague goals, shifting conditions and missing information. Naturalistic decision making (NDM) is a relatively contemporary approach which has emerged to account for decision making in such contexts (Klein, 1993; Zsombok & Klein, 1997) and has focused on decisions involving: ill-structured problems, uncertain dynamic environments, shifting or competing goals, time constraints. Certainly, police officers work in an environment which could be construed in this way and are required to make many complex decisions during the course of an investigation – including decisions concerning the validity, reliability or accuracy of CCTV evidence. Such decisions typically have high stakes, may involve risks and are accuracy motivated. To address the reality of real world decision

making, contemporary research adopting a NDM approach has considered critical decision making in firefighting (Weick, 1999), military operations (Kaempf, Klein, Thordsen, & Wolf, 1996), pilots in uncertain situations (Fischer, Orasanu et al., 1995), tactical operation commanders, (Schmitt & Klein, 1996) and doctors in emergency contexts (Crandall & Getchell-Reiter, 1993). Research in these areas has amply documented that experienced decision makers in these situations rarely make deliberated reasoned decisions or consider multiple alternatives but instead, make decisions based on their experiences and intuition (Crandall & Getchell-Reiter, 1993). They also rely on recognising parallels between a current context and previous similar context. In other words, they recognise an analogous situation based on available cues and then generate a plausible account based on experience – novices are typically unable to generate solutions in this manner (Orasanu & Connolly, 1993). Klein (1989) has argued that these recognition-primed decisions (RPDs) take place under time constraints, stress and changing or ambiguous conditions and involve “an assessment of the situation, recognition of events as typical, and a resultant course of action based on previous experience” (Randel & Pugh, 1996, p.580).

For instance, Klein (1998) found that experienced fire commanders did not appear to engage in a rational process of weighing up different options. Instead, if they recognised the type of situation, they intuitively knew the most likely successful course of action to take and as the course of action was successful on a previous occasion, there was no need to consider alternative solutions. Essentially, incoming information is evaluated with reference to what Klein, Phillips, Rall and Peluso (2007) call ‘frames’ (i.e. scripts, schemas or stories). Where inadequate, inaccurate or insufficient frames are adopted, the likelihood of poor decision making increases. For the purposes of CCTV review then, experienced officers need to be aware of the impact of prior experience on different cases may have on their reviewing decision – and use their experience to promote the accuracy of their review without letting it cloud their judgment.

## **7 *Variability in people***

### **7.1 Personality and other individual differences that affect performance**

It is important to start this section by pointing out that the research to be reviewed has almost always looked at the effect of one or, at most, two factors by which people differ. There is limited information about how personality factors and other individual differences interact. It is likely that the factors will work independently sometimes: if both improve performance individually, then the combination of factors improve it further. However, there will be no doubt other times where the interaction between factors is less predictable. Further, most of the factors investigated have at best a small impact on behaviour.

**Characteristics of people affecting understanding of degraded images.** An obvious starting point in understanding what about people may affect their ability to perceive images is that those with poor (uncorrected) visual acuity are likely to find degraded images difficult to process than those whose vision is normal or corrected to normal.

Identifying objects in degraded images to some degree requires people to separate the target from the scene it is in. This is one of the definitions used in the past for field independence, where field dependence/independence is considered a cognitive style that people adopt in many cognitive tasks (Witkin, 1964). Testing for it should be predictive of people's abilities to see what are essentially camouflaged objects. Typical tests include a Rod-and-Frame test and a Hidden Figures test. In the Rod-and-Frame test, people sit in the dark and see only an illuminated rod and frame. The frame is tilted, and the observer's goal is to align the rod to the horizon without being influenced by the frame. If they are successful, they are field independent. In the Hidden Figures test, people must pick out simple polygons that appear inside a drawn complex figure.

Earlier we reported studies showing that in relatively high quality videos, frame rate and whether the video was colour or B/W did not affect what information people got out of the videos (Chen et al, 2008; Gulliver et al, 2003). In the Chen et al study they also testing whether learning style (visual learner, verbal learner or bimodal) affected what could be picked up from the video versus the audio track, and whether image quality in these different dimensions would impact separately those with different learning styles. They found no effects that would suggest differential effects of image quality for those with different learning styles. The power of the study was low, in that not enough participants were in each condition, but the results showed no promising trends, either. In another study using the

same videos and assessing a similar question, cognitive style was assessed by means of a test of field dependence/independence (Ghinea & Chen, 2003). This, too, did not impact the effect of image quality on learning.

**Are some people better at identifying faces from surveillance footage than others?** To date, there is no evidence to suggest that people with experience of the identification process (such as experienced police officers) perform any better than untrained, inexperienced individuals when required to match or identify a previously unknown face. In Burton et al. (1999), police officers demonstrated the same low accuracy rates as student participants when attempting to identify and match an unfamiliar target face. Attempts to identify 'predictors' of face memory (or matching ability) have proved rather inconclusive. Schretlen, Pearlson, Anthony and Yates (2001) found a correlation between face matching ability, perceptual speed and cerebral volume while Alexander et al. (1999) reported individual differences in face matching performance were associated with activation of brain regions implicated in general object perception and attention system. More recently, Megreya and Burton (2006) examined this question in a series of six studies. Modest correlations between performance and test score were found for several standard tests of memory and perception (e.g. perceptual speed, visual short term memory and figure matching ability). However, the most interesting finding indicated that performance on a face matching task was predicted by performance on an inverted faces task (i.e. when faces were presented upside down). In other words, accuracy on the face matching task correlated with accuracy on an inverted faces task such that people who performed well on one task tended to also perform well on the other. People typically perform more poorly when attempting to identify inverted faces and evidence suggests that inverted faces are processed differently than faces presented in the usual upright position. Megreya and Burton (2006) concluded that the underlying reason for the poor accuracy rates typically observed on face matching tasks when the target is unfamiliar (e.g. when trying to decide whether an individual appearing in CCTV footage matches an image of a suspect) may be due to the fact that unfamiliar faces are processed in a different way to familiar faces.

**Characteristics of people affecting search.** Circadian rhythms affect search performance. Search for simple targets is up to 10% faster in the late afternoon than in the morning. However, search for more complex targets is faster in the early morning than in the afternoon (Monk, 1984). In a study of sleepiness and circadian rhythms in their effect on search, it was found that although responses were slower and less accurate at troughs in the circadian cycle, the direction of attention through the display was not (Horowitz, Cade, Wolfe

& Czeisler, 2003). It appears that what is affected by circadian rhythms is vigilance in making decisions about the information perceived, not how attention is focused in the displays.

**Characteristics of people affecting distraction/interruption.** When doing a review of CCTV that requires the reviewer to do complex event comprehension, the individual differences that cause variation in text comprehension can be assumed to apply. In particular, those who can keep more in mind comprehend text better (Daneman & Carpenter, 1983), and so ought to comprehend complex, unfolding events better. This has to do with the need to integrate information that is perceived earlier with information that is perceived later. Experts, when presented with a new scenario in their domain of expertise, immediately see and subsequently remember larger amounts of information than novices (Glaser & Chi, 1988; Ericsson & Kintsch, 1995). The memory skills of those with expertise in understanding behaviour, particularly criminal behaviour, would allow them to withstand interruptions better, as their memory would be more robust (Ledoux & Gordon, 2006; Oulasvirta & Saariluoma, 2006);

**Characteristics of people affecting vigilance.** There are notable individual differences in the relationship between time-on-task and vigilance. Craig (1984) reports that the variability between people accounts for more differences in vigilance performance than the variability between task and conditions, yet research has yet to be able to predict which individuals will do better than others.

There appears not to be any simple criterion available for what makes someone good at sustaining attention (Craig, 1984). Gender effects have been variable, and although aging may produce drops in raw performance, experience increases seem to offset this. Aptitude tests have not been able to reliably distinguish good from poor inspectors.

War veterans with post-traumatic stress disorder show poorer vigilance than war veterans with no symptoms of it (Vasterling, Brailey, Constans & Sutker, 1998). It is believed this happens because of being in an over-aroused state.

A number of personality characteristics have been tested for their relevance to the ability to maintain vigilance, many of which have been reviewed by Berch and Kanter (1984). Introverts are more sensitive to stimulation and extroverts are more impulsive, both of which seem relevant to vigilance, and so this personality factor has been tested a number of times. Early research found quite varied results. Other factors that have been tested include field dependence/independence, locus of control, Type-A vs Type B behaviour, and achievement

motivation. Field independent people, who are able to ignore irrelevant information better in looking for camouflaged objects, show better vigilance performance, as do people who believe the outcome of their behaviour is more under their own control than the control of outside forces, people with Type-A behaviour pattern, and people who are more sensitive to information about failures than information about successes. People who report higher levels of daydreaming during the day perform worse at vigilance tasks. Experienced meditators perform better than non-meditators and people who have recently begun to learn to meditate. All of these results are small enough that they would be difficult to use to select the best people to do a vigilance task.

Raby (2000, as cited by Washburn, Tagliatela, Rice & Smith, 2004) found that vigilance improved over time for some people and decreased for others. In an attempt to determine what skill might be predictive of this variability, Washburn et al asked people to do a simulated luggage-search task for 50 minutes, and on a separate day measured their ability to sustain attention in a different and simpler task. Those who performance in the top 25% on the sustained attention task also performed substantially better in the last 25 minutes of the luggage-search task than those who performed in the bottom 25% on the sustained attention task. The two groups of observers did not differ in their performance during the first 25 minutes of the luggage-search task. It should be noted that all observers were new to luggage search, and that their performance was improving rather than getting worse over time. Washburn et al assumed reasonably, but not definitively, that the smaller improvements in search for those with poorer sustained attention was due to poorer vigilance rather than poorer learning. Future research needs to tease apart vigilance and learning effects. It would also be important to test whether those experienced at doing luggage search would show the same relationship between their on-the-job performance and a simple sustained attention assessment.

Other studies, too, have looked at whether performance in different vigilance tasks is correlated, i.e., whether someone who is good at one vigilance task will be good at other vigilance tasks. The results have not been terribly promising, in that the more differences there are in the kind of vigilance task (e.g., simultaneous vs successive) the less consistency in performance is seen (Craig, 1984).

**Characteristics of people affecting change detection.** Younger adults are better than older adults on average, even when one controls for overall slowing due to age (Veiel, Storandt & Abrams, 2006). The youth advantage is seen not only in laboratory research, but

also when assessing information used to decide whether to turn at intersections (Caird, Edwards, Creaser & Horrey, 2005).

People who have strong imaging abilities generally do better at detecting changes than people who find it more difficult to form a mental image of what they have seen (Gur & Hilgard, 1975; Rodway, Gillies & Schepman, 2006).

Objects capture attention more when cognitive load is low (Lavie, 2005), and cognitive load tends to be lower in those with high working memory capacity. People with high working memory capacity also are better at understanding written text and speech, and in fact dealing with any information that needs to be integrated over time (Cowan, 2005). Therefore, there are a number of reasons to believe that those high in working memory capacity might perform better at CCTV review.

There is a measure used in a lot of applied research called functional field of view (FFOV), which measures how much of the world people will pick information up from without moving eyes or head (Ball, Roenker & Bruni, 1990). Those with a bigger FFOV perform better on change detection than those with a smaller one (Pringle, Irwin, Kramer & Atchley, 2001). FFOV tends to decrease with age, but FFOV impacted change detection both for younger adults and for older adults.

**Summary and Implications for CCTV review.** The literature on individual differences in the skills needed to effectively review CCTV footage is disjointed at best and probably not complete enough to strongly guide recommendations. Certainly there are tests to rule out truly disordered perception and thought processes and cognitive skills. But for a reasonably competent person, the remaining variability in perceptual and cognitive skills and in personality will not strongly predict performance on reviewing footage. More importantly, all the personality factors and cognitive skills have been tested in isolation rather than in combination, so the recommendations related to selection included in the Recommendations section of this report are at best tests to consider rather than tests for which there is firm evidence of relevance.

## 7.2 Training

**Training cognition: General.** A useful conceptualisation of what happens during training on a largely cognitive task is that first knowledge is acquired, then that knowledge is integrated into rules, and then those rules are internalised into mental procedures that require little



mental supervision (Companion & Gilson, 1988, as cited by Kass et al, 1991). When considering what differentiates good from poor reviewers of CCTV footage, then, it can be useful to consider whether they differ in terms of initial knowledge, organised use of that knowledge in their behaviour, or automatic prioritization of which rules are more important in which situations. Experts often see larger patterns or more meaningful information in the raw sensory data they perceive. Experts also tend to have better meta-cognitive skills, meaning they have a good sense about when they might be susceptible to errors and when they are not, they are better at planning multi-stepped cognitive behaviours, and can evaluate how efficient their behaviour is (Glaser & Chi, 1988).

When tested within their domain of expertise, experts appear to be able to hold more information in mind during the task, which has led to a theory that they develop a way to frame the information they see better so that it can be held longer (Ericsson & Kintsch, 1995). In addition, expert problem-solvers tend to find ways to restructure problems so that they are not as heavily dependent on maintaining information in working memory, a common limitation in cognitive performance. As mentioned before, naturalistic decision making research has shown that experienced decision makers tend to reason quickly and intuitively, and often by analogy, seeing the relationship between the situation they are in and other situations they have experienced before (Crandall & Getchell-Reiter, 1993; Orasanu & Connolly, 1993). Experts spend more time planning their behaviour than novices (Priest & Lindsay, 1992), and most of that time is spent in assessing the situation rather than laying out a course of action (Klein, Wolf, Militello, & Zsombok, 1995); this time spent planning results in less time doing the behaviour or correcting errors.

There seems to be ongoing planning by experts in tasks such as flying fighter aircrafts (Amaberti & Deblon, 1992), administering anaesthetics during surgery (Xiao, 1994, as cited by Dominguez, 2001) or performing surgery (Dominguez, 2001). Expert anaesthesiologists seem to be constantly anticipating situations that may occur in the near future and preparing how to respond to them (Xiao, 1994). Less experienced people, on the other hand, tend to behave more reactively than anticipatorily. Dominguez (2001) found anticipatory behaviour in expert laparoscopic surgeons, and in addition noted that, when asked to comment as they watched an operation the image the real surgeon saw about what they would be doing if they were the surgeon, the majority of experienced surgeons' comments about their perceptions also included a description of what those perceptions would mean for their actions during surgery, whereas interns made fewer such comments about the actions their perceptions afforded.

**Object recognition expertise.** There is no doubt that visual identification or classification/discrimination of objects can be improved by training. When people are given two sets of novel stimuli to discriminate from one another, practice improves the accuracy and speed of the judgments. However, the nature of the training is important. Discrimination of (odour) objects was better after training that required observers to identify the objects than after training that only required them to rate whether the objects had certain features, such as “sweetness” (Rabin, 1988). Identification training requires observers to discriminate between the objects they are judging. For example, if an observer were trying to decide whether any of a group of suspected criminals was amongst the people who passed through the scene a camera was recording, an observer who decided that a person seen was one of the suspects would be accompanied by fairly high confidence that that person was not one of the other suspects. On the other hand, if an observer were simply trying to identify people with brown hair, the observer could do the task without considering identify at all. Drawing trainees’ attention to the features of objects that are diagnostic is important: Identification of military aircraft after training was better when observers were told during training what features were most diagnostic of the different types of planes than when training presented objects without pointing out the distinctive features (Gibson, 1947, as reported by Proctor & Dutta, 1995). On difficult visual discrimination tasks, this can quickly raise the performance of novices to the level of experts (Biederman & Shiffrar, 1987).

The context in which targets to be identified are presented is critical to the effectiveness of training in generalizing identification skills to stimuli that have not yet been seen. Pellegrino, Doane, Fischer and Alderton (1991) trained people to discriminate target geometric shapes from non-target shapes. For some people, training used only nontargets that were fairly dissimilar to the target, and so judgments were easy to make. For other people, training used only nontargets that were fairly similar to the target, and so judgments were difficult. Training that focused on the most difficult discriminations generalised better to untrained comparisons than training that focused on the easiest discriminations. Proctor and Dutta (1995) describe this as due to a difference in representation that is built up during training: The simple-discrimination training set up a coarse representation of what the target was, which proved to be too imprecise to help with difficult discriminations, whereas the difficult-discrimination training set up a more precise representation of what a target was, sufficient for all later discriminations. Later research showed that there was more to what was learned than the representation of the specific target used in search. When people were transferred to making discrimination judgments about completely different target and non-target polygons, those who made the more difficult discriminations in training still outperformed

those who made the easier discriminations in training (Doane, Alderton, Sohn & Pellegrino, 1996).

A later study showed that training difficulty did not have the same effect for all observers (Sohn, Doane & Garrison, 2006). When cognitive ability of the observers was assessed, those with mid-range abilities showed the largest effect of training difficulty. Those with high ability learned well regardless of training, and those with low ability did not learn well regardless of training.

The broader context of the perceptual world in which target objects are trained also influences what is learned. One of the skills that people must learn is which information in the perceptual world is relevant to identifying targets. Kass, Herschler and Companion (1991) asked university students to identify the direction from which a (simulated) military tank was approaching. The relevant information was the pattern of asymmetry in the muzzle flash produced by the tank's firing. The full task during testing included much other visual information from the simulated battlefield, as well as accompanying auditory information. They compared training which only presented the muzzle flashes to be judged, with no other visual or auditory scene, to training in which the tank flashes were presented in the context of the full scene. Note that some of this information included redundant cues (albeit weak ones) about the approach of the tank. They found the simpler training generalised better than the more realistic training, even though both groups received initial training about how to interpret the pattern of the muzzle flashes.

**Object recognition expertise, degraded stimuli.** Fiore, Scielzo and Jentsch (2004) ran a study similar to that of Kass et al (1991) using X-ray images of luggage like those produced for carry-on luggage in airports. People were either trained to identify guns and knives that appeared alone within a piece of luggage or were trained to identify them when they appeared amid other objects, where the multiple objects did not overlap. Interestingly, efficacy of the training was not straightforwardly predicted by the nature of the training trials. Instead, observers who came to the task with strong spatial skills benefited most by full-luggage training, whereas observers who came to the task with weaker spatial skills benefited most by isolated-object training. Thus, training style interacted with cognitive skills. In assessing how this relates to training of CCTV footage review, it is important to note that the particular skill involved in luggage search, spatial skill, may not be critically relevant in that it is a skill in mentally rotating what is seen in one orientation to a different orientation that was seen before. This is very critical to identifying objects in luggage, but is probably less important to identifying objects in CCTV footage. Any training developed for CCTV

footage review should identify the most important cognitive skills, measure individual differences in those skills, and test how training in using degraded images interacts with those cognitive skills.

What cognitive skills are likely to be important? Integrating information from multiple perspectives or views, detecting changes from one view to the next, understanding intentional behaviour all seem likely candidates.

**Training search.** Search gets better with practice. If targets are not obvious, search accuracy improves along with search speed. When targets are obvious and so accuracy has no room for improvement, speed of search improves. It is important to note that practice effects can often override vigilance decrements, resulting in improving performance even when vigilance is waning (Washburn et al, 1984).

**Training vigilance.** Although there is little support that one can select people for their vigilance abilities, training has been shown to improve performance, at least in the short term (Craig, 1984). Feedback about performance helps observers to maintain vigilance, but only if the monitoring task maintains its simultaneous or successive characteristic rather than switching from one to the other (Becker, Warm, Dember, & Howe, 1994). Thus, formal training at CCTV footage review would be beneficial as long as the training provided feedback about whether the reviewer successfully found targets and provided the training matched the simultaneous/successive nature of the reviewing to be done after training. On-the-job training is less likely to be successful unless an experienced reviewer monitors the novice's vigilance and provides feedback.

**Training change detection.** It is not common for change blindness to improve within the course of a laboratory experiment, which leads some people to believe that change detection is impervious to training. However, there is other research showing that people who are experts in the domain in which they are looking for changes reliably detect more changes than those who are novices in the domain (Werner & Thies, 2000); these changes only apply when looking at scenes relevant to the domain of expertise. This no doubt relates to the finding that changes to objects of central interest in a scene's meaning will be detected better than changes to objects of peripheral interest (Rensink, O'Regan, & Clark, 2000). Experts will understand the deep meaning of a scenario better than novices (Glaser & Chi, 1988), and so will be more sensitive to changes related to that meaning. This suggests that one way to improve change detection in CCTV footage is to give general training on what the meaning of the footage is. Footage varies widely, but those with greater policing

experience and training will generally be better at assessing what is going on in behaviour being observed.

**Training decision making.** Decision making does improve with practice, but the nature of the training needed depends on the domain in which decisions are made. Klein (1997) identifies the following as relevant to address if one wants to improve decision making skills:

- Gaining awareness of one's decision making process, which can be aided by sampling alternative decision-making practices
- Practicing the real task, as expertise always develops because of experience
- Practice attention shifting, as situational awareness depends on attention
- Getting feedback from others about one's performance and reflecting on one's experiences

**Summary and Implications for CCTV review.** Processing of degraded stimuli should improve with practice, particularly if there is training about diagnostic features of different kinds of target objects or events and about strategies to deal with the degraded stimuli. Training should give feedback about accuracy in reviewing the footage and should be tailored to the skills the reviewer brings to the task, as people with different cognitive skills and abilities benefit from different kinds of training.

Increasing experience will improve the expectations CCTV reviewers build throughout the review process based on prior knowledge and on what they are seeing. In many cases, experience-based expectations may be useful and guide accurate decision-making while in other situations, such expectations may be a hindrance as they will render reviewers more vulnerable to rigidity in their search strategy and situation interpretation. Thus, awareness of the potential problems is important to help reviewers maintain their cognitive flexibility. Explicit training early on and periodic reminders about the dangers of adopting the wrong mindset and training on maintaining cognitive flexibility should lower the likelihood of there being problems.

## **8 Report Conclusions**

The review of CCTV footage requires many cognitive skills, and those skills need to be learned, coordinated and maintained. Optimise performance in reviewing footage presents challenges that have not yet been directly researched. We have outlined a number of recommendations for optimising performance, based on current understanding of the skills involved, but there is much that could be learned by targeted research.

## References

- Alexander, G. E., Mentis, M. J., Van Horn, J. D., Grady, C. L., Berman, K. F., Furey, M. L. (1999). Individual differences in PET activation of object perception and attention systems predict face matching accuracy. *NeuroReport*, 10, 1965-1971.
- Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation based model. *Cognitive Science*, 26, 39-83.
- Anderson, C., & Home, J. A. (2006). Sleepiness enhances distraction during a monotonous task. *Sleep: Journal of Sleep and Sleep Disorders Research*, 29(4), 573-576.
- Anderson, R. A., & Bradley, D. C. (1998). Perception of three-dimensional structure from motion. *Trends in Cognitive Sciences*, 2(6), 222-228.
- Apteker, R. T., Fisher, J. A., Kisimov, V. S., & Neishlos, H. (1995). Video acceptability and frame rate. *IEEE Multimedia*, 2, 32-40.
- Ask, K., & Granhag, P. A. (2005). Motivational sources of confirmation bias in criminal investigations: The need for cognitive closure. *Journal of Investigative Psychology and Offender Profiling*, 2, 43-63.
- Ask, K., & Granhag, P. A. (2007). Motivational bias in criminal investigators' judgments of witness reliability. *Journal of Applied Social Psychology*, 37, 561-591.
- Ask, K., Rebelius, A. & Granhag, P-A. (2008). The 'elasticity' of criminal evidence: A moderator of investigator bias. *Applied Cognitive Psychology* (Online: <http://www3.interscience.wiley.com/cgi-bin/fulltext/117878349/PDFSTART>)
- Avrahami, J., & Kareev, Y. (1994). The emergence of events. *Cognition*, 53, 239-261.
- Baldwin, J. (1993). Police interview techniques; establishing truth of proof. *British Journal of Criminology*, 33, 325-352.
- Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants detect structure in human action: A first step toward understanding others' intentions? *Child Development*, 72, 708- 718.
- Becker, A. B., Warm, J. S., Dember, W. N., & Howe, S. (1994). Specific and non-specific transfer effects in training for vigilance. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp. 294-299). Hillsdale, NJ: Erlbaum.
- Bell, R., Cardello, A. V., & Schutz, H. G. (2003). Relations among comfort of fabrics, ratings of comfort, and visual vigilance. *Perceptual and Motor Skills*, 97(1), 57-67.
- Benton, A. L. (1980). The neuropsychological of facial recognition. *American Psychologist*, 35, 176-186.

- Berch, D. B., & Kanter, D. R. (1984). Individual differences. In Warm, J. S. (Ed.) *Sustained Attention in Human Performance* (pp. 143-178). New York: John Wiley & Sons.
- Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 640-645.
- Bodenhause, G. V. (1990). Second-guessing the jury: Stereotypic and hindsight biases in perceptions of court cases. *Journal of Applied Social Psychology*, 20, 1112-1121.
- Boschman, M. C., Leermakers, M. A., & Roufs, J. A. (1985). The effect of video bandwidth on the judgment of comfort, visual performance and eye movements using visual display units. *IPO Annual Progress Report*, 20, 73-79.
- Boschman, M. C., & Roufs, J. A. (1989). The effect of bandwidth and luminance contrast of VDUs and print on visual comfort and performance. *IPO Annual Progress Report*, 24, 72-80.
- Boschman, M. C., & Roufs, J. A. (1997). Text quality metrics for visual display units: II. An experimental survey. *Displays*, 18(1), 45-64.
- Bradley, B. P., Mogg, K., Wright, T., & Field, M. (2003). Attentional bias in drug dependence: Vigilance for cigarette-related cues in smokers. *Psychology of Addictive Behaviors*, 17, 66-72.
- Bruce, V., Henderson, Z., Greenwood, K., Hancock, P., Burton, A.M. & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, 5, 339-360.
- Bruce, V., Henderson, Z., Newman, C. & Burton, A.M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, 7, 207-218.
- Bruner, J. S., & Potter, M. C. (1964). Interference in visual recognition. *Science*, 144, 424-425.
- Burton, A.M., Miller, P., Bruce, V., Hancock, P.J.B. & Henderson, Z. (2001). Human and automatic face recognition: a comparison across image formats. *Vision Research*, 41, 3185-3195.
- Burton, A.M, Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor quality video: evidence from security surveillance. *Psychological Science*, 10, 243-248.
- Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement. *Psychonomic Bulletin & Review*, 11, 932-937.
- Caird, J. K., Edwards, C., Creaser, J., & Horrey, W. J. (2005). Older driver failures of attention at intersections: Using change blindness methods to assess turn decision accuracy. *Human Factors*, 47, 235-249.

- Chen, S. Y., Ghinea, G., & Macredie, R. D. (2006). A cognitive approach to user perception of multimedia quality: An empirical investigation. *International Journal of Human-Computer Studies*, 64, 1200-1213.
- Cheung, S. S., Westwood, D. A., & Knöz, M. K. (2007). Mild body cooling impairs attention via distraction from skin cooling. *Ergonomics*, 50(2), 275-288.
- Cohen, M. S. (1989). A database tool to support probabilistic assumption-based reasoning in intelligence analysis. *Proceedings of the 1989 Joint Director of the C2 Symposium*, Ft. McNair, VA, June.
- Cohen, N., Gattuso, J., & MacLennan-Brown, K. (2006). *CCTV Operational Requirements Manual, Version 4.0*. Publication 55/06. Home Office Scientific Development Branch. Retrieved from [www.hosdb.homeoffice.gov.uk](http://www.hosdb.homeoffice.gov.uk).
- Companion, M. A., & Gilson, R. D. (1988). *Training technology for situational awareness*. Unpublished manuscript.
- Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics*, 58, 602-612.
- Costigan, R. (2007). Identification from CCTV: the risk of injustice. *Criminal Law Review*, 591-608.
- Cowan, N. (2005). *Working memory capacity*. New York, NY: Psychology Press.
- Craig, A. (1984). Human engineering: The control of vigilance. In Warm, J. S. (Ed.) *Sustained Attention in Human Performance* (pp. 247-292). New York: John Wiley & Sons.
- Crandall, B., Klein, G. & Hoffman, R. R. (2006). *Working Minds: A Practitioner's Guide to Cognitive Task Analysis*. London: MIT Press.
- Crandall, B., & Getchell-Reiter, K. (1993) Critical Decision Method: A technique for eliciting concrete assessment indicators from the intuition of NICU nurses. *Advances in Nursing Science*. 16, 42-51.
- Cranley, N., Murphy, L., Perry, P. (2003). User-perceived quality-aware adaptive delivery of MPEG-4 content. *Proceedings of the 13<sup>th</sup> International Workshop on Network and Operating Systems Support for Digital Audio and Video*. Monterey, CA. pp. 42-49.
- Daneman, M., & Carpenter, P. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(4), 561-584.
- Darley, J. M., & Gross, P. H. (1983). A hypothesis confirming bias in labelling effects. *Journal of Personality and Social Psychology*, 44, 20-33.
- Davies, D. R., & Parasuraman, R. (1982). *The Psychology of Vigilance*. London: Academic Press.



- Davies, G., & Hine, S. (2007). Change blindness and eyewitness testimony. *The Journal of Psychology*, 141(4), 423-434.
- Davies, G. & Thasen, S. (2000). Closed circuit television: how effective an identification air? *British Journal of Psychology*, 91, 411-26.
- Dawes, R. M. (1988). *Rational choice in an uncertain world*. New York: Harcourt Brace Jovanovich.
- Dijker, A. J. & Koomen, W. (1996). Stereotyping and attitudinal effects under time pressure. *European Journal of Social Psychology*, 26, 61-74.
- Ditto, P. H., Munro, G. D., Apanovitch, A. M., Scepansky, J. A., & Lockhart, L. K. (2003). Spontaneous skepticism: The interplay of motivation and expectation in responses to favorable and unfavourable medical diagnoses. *Personality and Social Psychology Bulletin*, 29, 1120–1132.
- Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology*, 75, 53–69.
- Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception*, 22, 15-22.
- DiVita, J., Obermeyer, R., Nygren, T. E., & Linville, J. M. (2004). Verification of the change blindness phenomenon while managing critical events on a combat information display. *Human Factors*, 46, 205-218.
- Doane, S. M., Alderton, D. L., Sohn, Y. W., & Pellegrino, J. W. (1996). Achisition and transfer of skilled performance: Are visual discrimination skills stimulus specific? *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1218-1248.
- Dror, I. E., Busemeyer, J. R. & Basola, B. (1999). Decision Making Under Time Pressure: An independent test of sequential sampling models. *Memory and Cognition*, 27, 713-725.
- Dror, I. E. & Charlton, D. (2006). Why experts make errors. *Journal of Forensic Identification*, 56, 600-616.
- Dror, I. E., Charlton, D., & Peron A. (2006). Contextual information renders experts vulnerable to making erroneous identifications *Forensic Science International*, 156, 74-78.
- Dror, I. E., Peron, A., Hind, S., & Charlton, D. (2005). When emotions get the better of us: The effect of contextual top-down processing on matching fingerprints. *Applied Cognitive Psychology*, 19,799-809.
- Dror, I. E. & Rosenthal, R. (2008). Meta-analytically quantifying the reliability and biasability of forensic experts. Manuscript in press at *Journal of Forensic Sciences*, 53.

- Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology*, 34, 590-598.
- Durlach, P. J. (2004). Change blindness and its implications for complex monitoring and control systems design and operator training. *Human Computer Interaction*, 19, 423-451.
- Dominguez, C. O. (2001) Expertise in laparoscopic surgery: Anticipation and affordances. In: E. Salas and G. Klein, Editors, *Linking Expertise and Naturalistic Decision Making*, 287-301. Mahwah, NJ: Lawrence Erlbaum Associates.
- Edland, A., & Svenson, O. (1993). Judgment and decision making under time pressure: Studies and findings. In O. Svensson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 27-40). New York: Plenum Press.
- Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology*, 71, 5-24
- Edwards, M. B., & Gronlund, S. D. (1998). Task interruption and its effects on memory. *Memory*, 6, 665-687.
- Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology*. 48, 269-297.
- Elstein, A. S., Schulman, L. S., & Sprafka, S. A. (1978). *Medical Problem Solving: An Analysis of Clinical Reasoning*. Cambridge, Mass: Harvard University Press.
- Erickson, R. A. (1966). Visual search experiment: Noise persistence, acuity, response time. *Journal of the Optical Society of America*, 56, 491-498.
- Fiore, S. M., Scielzo, S., & Jentsch, F. (2004). Stimulus competition during perceptual learning: Training and aptitude considerations in the X-ray security screening process. *International Journal of Cognitive Technology*, 9(2), 34-39.
- Fiorentini, A., Maffei, L., & Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception*, 12, 195-201.
- Fischer, U., Orasanu, J. M. & Wich, M. (1995) Expert pilots' perception of problem situations. In R. Jensen (Ed.). *Proceedings of the Eighth International Symposium on Aviation Psychology*. April 24-27, 1995, Columbus, OH.
- Fisk, A. D., & Schneider, W. (1981). Controlled and automatic processing during tasks requiring sustained attention. *Human Factors*, 23, 737-750.
- Folk, C. L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception & Performance*, 24(3), 847-58.
- Foster, G., Schmidt, C., & Sabatino, D. (1976). Teacher expectancies and the label "learning disabilities." *Journal of Learning Disabilities*, 9, 111-114.

- Fukuda, K., Wakamiya, N., Murata, M., & Miyahara, H. (1997). QoS Mapping between user's preference and bandwidth control for video transport. *Proceedings of the Fifth International Workshop on QoS (IWQoS)*, New York.
- Ghinea, G., & Chen, S.-Y. (2003). The impact of cognitive styles on perceptual distributed multimedia quality. *British Journal of Educational Technology*, 34(4), 393–406.
- Gibson, J. J. (Ed.). (1947). *Motion picture testing and research* (Army Air Forces Aviation Psychology Program Research Reports, Report No. 7). Washington, D.C.: U.S. Government Printing Office.
- Glanzer, M., Fischer, B., Dorfman, D., 1984. Short-term storage in reading. *Journal of Verbal Learning and Verbal Behavior* 23, 467-486.
- Glaser, R., & Chi, M. T. H. (1988). Overview. In M.T.H. Chi, R. Glaser, and M. J. Farr (Eds.) *The Nature of Expertise* (pp. xv-xxvii). Hillsdale, NJ: Erlbaum.
- Gulliver, S. R., Serif, T., & Ghinea, G. (2004). Pervasive and standalone computing: the perceptual effects of variable multimedia quality. *International Journal of Human-Computer Studies*, 60, 640-665.
- Gulliver, S. R., & Ghinea, G. (2004). Stars in their eyes: What eye-tracking reveals about multimedia perceptual quality. *IEEE Transaction on System, Man and Cybernetics, Part A*, 34(4), pp 472-482.
- Gur, R. C., & Hilgard, E. R. (1975). Visual imagery and the discrimination of differences between altered pictures simultaneously and successively presented. *British Journal of Psychology*, 66, 341-345.
- Haines, R. F. (1991). A breakdown in simultaneous information processing. In G. Obrecht & L. Stark (Eds.), *Presbyopia research: From molecular biology to visual adaptation* (pp. 171-175). New York: Plenum.
- Harville, D. L., Lopez, N., Elliott, L. R., & Barnes, C. (2005). *Team communication and performance during sustained working conditions*. Interim report AFRL-HE-BR-TR-2005-0085 of the Air Force Research Labs.
- Helton, W. S., Hollander, T. D., Warm, J. S., Matthews, G., Dember, W. N., Wallaart, M., Beauchamp, G., Parasuraman, R., & Hancock, P. A. (2005). Signal regularity and the mindlessness model of vigilance. *British Journal of Psychology*, 96, 249-261.
- Henderson, J. M., Brockmole, J. R., Castelano, 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 movements: A window on mind and brain* (pp. 537-562). Oxford: Elsevier.
- Henderson, Z., Bruce, V. & Burton, A.M. (2001). Matching the faces of robbers captured on video. *Applied Cognitive Psychology*, 15, 445-464.

- Herman, L. M., Morrel, S. P. & Pack, A. A. (1990). Bottlenosed dolphin and human recognition of veridical and degraded video displays of an artificial gestural language. *Journal of Experimental Psychology: General*, 119, 215-230
- Hesse, L. (2002). The transition from video motion detection to intelligent scene discrimination and target tracking in automated video surveillance systems. *Security Journal*, 15, 69-78.
- Hill, H. & Bruce, V. (1996). Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 986-1004.
- Hill, H., & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces'. *Current Biology*, 11, 880-885.
- Hillstrom, A. P., & Chai, Y. C. (2006). Factors that guide or disrupt attentive visual processing. *Computers in Human Behavior*, 22, 648-656.
- Hillstrom, A. P., & Yantis, S. (1994). Visual motion and attention capture. *Perception & Psychophysics*, 20(1), 95-107.
- Hikichi, K., Morino, H., Matsumoto, S., Yasuda, Y., Arimoto, I., Ijume, M., & Sezaki, K. (2001). Architecture of Haptics communication system for adaptation to network environments. *Proceedings of the IEEE International Conference on Multimedia and Expo*. Tokyo, Japa. Pp. 744-747.
- Hirst, E., & Schweitzer, M. (1990). Electric-utility resource planning and decision-making: The importance of uncertainty. *Risk Analysis*, 10, 137-146.
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, 24, 1-55.
- Hollingworth, A., & Henderson, J. M. (2003). Testing a conceptual locus for the inconsistent object change detection advantage in real-world scenes. *Memory & Cognition*, 31, 930-940.
- Hollingworth, A., & Henderson, J. M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Visual Cognition*, 7, 213-235.
- Horowitz, T. S., Cade, B. E., Wolfe, J. M., & Gzeisler, C. A. (2003) Searching night and day: a dissociation of effects of circadian phase and time awake on visual selective attention and vigilance. *Psychological Science*, 14(6), 549-557.
- Hope, L., Memon, A. & McGeorge, P. (2004) Understanding pre-trial publicity: Pre-decisional distortion of evidence by mock jurors. *Journal of Experimental Psychology: Applied*, 10, 111-119.
- Itti, L. Koch, C. & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1254-1259.
- J. (2001). Cutting Edge of Technology: The Use of CCTV/Video Cameras in Law Enforcement.

- Janis, I. L., & Mann, L. (1977). *Decision-making: A psychological analysis of conflict, choice and commitment*. New York: Free Press.
- Jastorff, J., Kourtzi, Z., & Giese, M. A. (2006). Learning to discriminate complex movements: Biological versus artificial trajectories. *Journal of Vision*, 6(8), 791-804.
- Jokisch, D., Daum, I., & Troje, N. F. (2006). Self recognition versus recognition of others by biological motion: Viewpoint-dependent effects. *Perception*, 35, 911-920.
- Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology*, 80, 557-571.
- Kaempff, G.L., Klein, G.A., Thordsen, M.L. & Wolf, S. (1996). Decision making in complex naval command-and-control environments. *Human Factors*, 38, 220-231.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kalven, H., & Zeisel, H. (1966). *The American Jury*. Boston: Little Brown.
- Kass, S. J., Herschler, D. A., & Companion, M. A. (1991). Training situational awareness through pattern recognition in a battlefield environment. *Military Psychology*, 3(2), 105-112.
- Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing: photographs, credit cards and fraud. *Applied Cognitive Psychology*, 11, 211-222.
- Kennedy, A., Brysbaert, M., & Murray, W. S. (1998). The effects of intermittent illumination on a visual inspection task. *The Quarterly Journal of Experimental Psychology*, 51A(1), 135-151.
- Kennedy, A., & Murray, W. S. (1991). The effects of flicker on eye movement control. *Quarterly Journal of Experimental Psychology*, 43A, 79-99.
- Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review*, 95, 163-182.
- Kjellberg, A. (1990). Subjective, behavioural and psychophysiological effects of noise. *Scandinavian Journal of Work and Environmental Health*, 16 (sup 1), 29-38.
- Klein, G. A. (1989). Recognition-primed decisions. In W. B. Rouse, Ed. *Advances in Man-Machine Systems Research*, 5, 47-92.
- Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision-making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. Zsombok (Eds.), *Decision-Making in Action: Models and Methods*, (138-147). Norwood, CT: Ablex.
- Klein, G. A. (1997). Developing expertise in decision making. *Thinking and Reasoning*, 3, 337-352.

- Klein, G. A. (1998). *Sources of Power: How People Make Decisions*. Cambridge, MA: MIT Press.
- Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. In R. R. Hoffman (Ed.), *Expertise out of context* (pp. 113–155). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Klein, G., Wolf, S., Militello, L., & Zsombok, C. (1995). Characteristics of skilled option generation in chess. *Organizational Behavior and Human Decision Processes*, 62, 63-69.
- Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, 110, 499–519.
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 107-118.
- Kourtzi, Z., & Nakayama, K. (2002). Distinct mechanisms for the representation of moving and static objects. *Visual Cognition*, 9, 248-264.
- Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay inferences: Effects of impression primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, 19, 448-468.
- Koelega H. S. (1993). Stimulant drugs and vigilance performance: a review. *Psychopharmacology*, 111, 1-16.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 451-468.
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, 9, 75-82.
- Law, A. S., Logie, R. H., Pearson, D. G., Cantagallo, A., Moretti, E., & Dimarco, F. (2004). Resistance to the impact of interruptions during multitasking by healthy adults and dysexecutive patients. *Acta Psychologica*, 116, 285-307.
- Ledoux, K., & Gordon, P. C. (2006). Interruption-similarity effects during discourse processing. *Memory*, 14, 789-803.
- Leermakers, M.A.M., & Boschman, M. C. (1984). Eye movements, performance and visual comfort using VDTs. *IPO Annual Progress Report*, 19, 70-75.
- Leo, R. A. (1996). Inside the interrogation room. *Journal of Criminal Law and Criminology*, 86, 266-303.
- Levin, DD. T., Momen, N., Drivdahl, S. B., & Simons, D. J. (2000). Change blindness blindness: The metacognitive error of overestimating change-detection ability. *Visual Cognition*, 7, 397-412.
- Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, 4, 501-506.

- Lieberman, H. R. (2003). Nutrition, brain function and cognitive performance. *Appetite*, 40(3), 245-254.
- Lieberman, H. R., Kramer, F. M., Montain, S. J., & Niro, P. (2007). Field assessment and enhancement of cognitive performance: Development of an ambulatory vigilance monitor. *Aviation, Space, and Environmental Medicine*, 78, B268-B275.
- Lipshitz, R., & Ben Shaul, O. (1997). Schemata and mental models in recognition-primed decision-making. In C. Zsombok, & G. A. Klein (Eds.), *Naturalistic decision-making* (pp. 292–303). Hillsdale, NJ: Erlbaum.
- Lleras, A., Rensink, R. A., & Enns, J. T. (2005). Rapid resumption of interrupted visual search. *Psychological Science*, 16, 684-688.
- Lleras, A., Rensink, R. A., & Enns, J. T. (2007). Consequences of display changes during interrupted visual search: Rapid resumption is target specific. *Perception & Psychophysics*, 69, 980-993.
- Logie, R. H., Baddeley, A. D., & Woodhead, M. M. (1987) Face recognition, pose and ecological validity. *Applied Cognitive Psychology*, 1, 53-69.
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37, 2098–2109.
- Lundgren, S. R. & Prislin, R. (1998). Motivated cognitive processing and attitude change. *Personality and Social Psychology Bulletin*, 24, 715-26.
- Mack, A., & Rock, I. (1998). *Inattention blindness*. Cambridge, MA: MIT Press.
- Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, 1, 6-21.
- Marshall, S. P. (2007). Identifying cognitive state from eye metrics. *Aviation, Space and Environmental Medicine*, 78, B165-B175.
- Megreya, A. M. & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition*, 34, 865-876.
- Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: Effects on eyewitness accuracy and confidence. *British Journal of Psychology*, 94, 339-354.
- Metropolitan Police (2005) 'Police Investigation Continues into the 7/7 Bombings' *Metropolitan Police Bulletin* 222, 18 July 2005. Available from: <[http://cms.met.police.uk/news/major\\_operational\\_announcements/terrorist\\_attacks/police\\_investigation\\_continues\\_into\\_the\\_7\\_7\\_bombings](http://cms.met.police.uk/news/major_operational_announcements/terrorist_attacks/police_investigation_continues_into_the_7_7_bombings)>. Downloaded: 12/03/08.
- Mocharnuk, J. B., Gaudio, D. F., & Suwe, C. L. (1981). Imaging infrared ship target acquisition studies. *Human Factors*, 23, 561-580.
- Monk, T. H. (1984). Search. In Warm, J. S. (Ed.) *Sustained Attention in Human Performance* (pp. 293-322). New York: John Wiley & Sons.



- Most, S. B., Scholl, B. J., Clifford, E., & Simons, D. J. (2005). What you see is what you set: Sustained inattention blindness and the capture of awareness. *Psychological Review*, 112, 217-242.
- Nachreiner, F. (1977). Experiments on the validity of vigilance experiments. In R. R. Mackie (Ed.) *Vigilance: Theory, Operational Performance and Physiological Correlates*, pp 665-678. New York: Plenum.
- NACRO (2002). *To CCTV or Not to CCTV? A Review of Current Research into the Effectiveness of CCTV Systems in Reducing Crime*. London, NACRO. Publication Number 2002062800.
- Näsänen, R., & Ojanpää, H. (2003). Effect of image contrast and sharpness on visual search for computer icons. *Displays*, 24, 137-144.
- Näsänen, R., Karlsson, J., & Ojanpää, H. (2001). Display quality and the speed of visual letter search. *Displays*, 22, 107-113.
- Newton, D., Engquist, G., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology*, 35, 847-862.
- Nguyen, L. T., Jauregui, B., & Dinges, D. F. (1998). *Changing behaviours to prevent drowsy driving and promote traffic safety: Review of proven, promising and unproven techniques*. Report to the American Automobile Association Foundation for Traffic Safety. [www.aaafoundation.org/pdf/drowsydriving.pdf](http://www.aaafoundation.org/pdf/drowsydriving.pdf)
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2, 175-220.
- Nisbett, R. E., & Ross, L. D. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Oakhill, J., Garnham, A. & Reynolds, D. (2005). Immediate activation of stereotypical gender information, *Memory & Cognition*, 33, 972-983.
- Orasanu, J. and Connolly, T. (1993) The Reinvention of Decision Making, in Klein, G. A., Orasanu, J., Calderwood, R. and Zsombok, C. E. (eds.) *Decision Making in Action: Models and Methods*, pp 3-20, Ablex, Norwood, NJ.
- Oulasvirta, A., & Saariluoma, P. (2006). Surviving task interruptions: Investigating the implications of long-term working memory theory. *International Journal of Human-Computer Studies*, 64, 941-961.
- Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology & Behavior*, 93(1-2), 369-378.
- Parasuraman, R., & Davies, D. R. (1977). A taxonomic analysis of vigilance performance. In R. R. Mackie (Ed.), *Vigilance: Theory, Operational Performance, and Physiological Correlates* (pp. 559-574). New York: Plenum Press.



- Parrott, A. C., Garnham, N. J., (1996). Cigarette smoking and abstinence: Comparative effects upon cognitive task performance and mood state over 24 hours. *Human Psychopharmacology: Clinical and Experimental*, 11, 391-400.
- Pearson, J. E., & Pearson, C. E. (1985). The effects of bandwidth compression and image quality on image interpreter performance, *Human Factors*, 27, 345-353.
- Peissig, J. J., & Tarr, M. J. (2007). Visual object recognition: Do we know more now than we did 20 years ago? *Annual Review of Psychology*, 58, 75-96.
- Pellegrino, J. W., Doane, S. M., Fischer, S. C., & Alderton, D. (1991). Stimulus complexity effects in visual comparisons: The effects of practice and learning context. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 781-791.
- Pennington, N. & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of Personality & Social Psychology*, 51, 242-258.
- Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 521-533.
- Pennington, N. & Hastie R. (1993). A theory of explanation-based decision making. In G.A. Klein, J. Orasanu, J. Caldentood and D. MacGregor (eds) *Decision Making in Action: Models and Methods*. Norwood NJ, Ablex Publishing.
- Perry, A. R., & Laurie, C. A. (1992). Sustained attention and the Type A behavior pattern: the effect of daydreaming on performance. *Journal of General Psychology*, 119, 217-228.
- Pigeau, R.A., Angus, R.G., O'Neill, P., & Mack, I. (1995). Vigilance latencies to aircraft detection among NORAD surveillance operators. *Human Factors*, 37, 622-634.
- Pilcher, J. J., Nadler, E., & Busch, C. (2002). Effects of hot and cold temperature on performance: A meta-analytic review. *Ergonomics*, 45(10), 682-698.
- Priest, A. G., & Lindsay, R. O. (1992). New light on novice-expert differences in physics problem solving. *British Journal of Psychology*, 83(3), 389-405.
- Pringle, H. L., Irwin, D. E., Kramer, A., F., & Atchley, P. (2001). The role of attentional breadth in perceptual change detection. *Psychonomic Bulletin & Review*, 8, 89-95.
- Proctor, R. W., & Dutta, A. (1995). *Skill Acquisition and Human Performance*. London: Sage Publications.
- Pyszczynski, T., & Greenberg, J. (1987). Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. *Advances in Experimental Social Psychology*, 20, 297-340.
- Rabin, M. D. (1988). Experience facilitates olfactory quality discrimination. *Perception & Psychophysics*, 44, 532-540.
- Rajaram, S. (1998). The effects of conceptual salience and perceptual distinctiveness on conscious recollection. *Psychonomic Bulletin & Review*, 5, 71-78.

- Randel, J. M. & Pugh, H. L. (1996). Differences in expert and novice situation awareness in naturalistic decision making. *International Journal of Human-Computer Studies*, 45, 579-597.
- Remington, R., Folk, C. L., & McLean, J. P. (2001). Contingent attentional capture or delayed allocation of attention? *Perception & Psychophysics*, 63 (2), 298-307.
- Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes *Psychological Science*, 8, 368-373.
- Rensink, R. A., O'Regan, J. K., & Clark, J. J. (2000). On the failure to detect changes in scenes across brief interruptions. *Visual Cognition*, 7, 127-145.
- Roark, D. A., O'Toole, A. J., Abdi, H., & Barrett, S. E. (2006). Learning the moves: The effect of familiarity and facial motion on person recognition across large changes in viewing format. *Perception*, 35(6), 761-773.
- Robertson, I. H., Manly, T., Andrade, T., Baddeley, B. T., & Yiend, J. Y. (1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, 24, 636-647.
- Rodway, P., Gillies, K., Schepman, A. (2006). Vivid imagers are better at detecting salient changes. *Journal of Individual Differences*, 27, 218-228.
- Santangelo, V., & Spence, C. (in press). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology: Human Perception and Performance*.
- Saylor, M. M., & Baldwin, D. A. (2004). Action analysis and change blindness: Possible links. In D. T. Levin (Ed.) *Thinking and seeing: Visual metacognition in adults and children*. (pp. 37-56). Cambridge, MA: MIT Press.
- Schmitt, J. F. & Klein, G. A. (1996). Fighting in the fog: Dealing with battlefield uncertainty. *Marine Corps Gazette*, 80, 63-66.
- Scholl, B. J. (2000). Attenuated change blindness for exogenously attended items in a flicker paradigm. *Visual Cognition*, 7, 377 - 396.
- Schretlen, D. J., Pearlson, G. D., Anthony, J. C., Yates, K. O. (2001). Determinants of Benton Facial Recognition Test performance in normal adults. *Neuropsychology*, 15, 405-410.
- Schulz-Hardt, S., Frey, D., Lüthgens, C., & Moscovici, S. (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology*, 78, 655-669.
- Scott, D. (1996). A pilot study of nonverbal cues in videotelecommunication. *Psychological Reports*, 78, 555-561.
- Scott-Brown, K. C. & Cronin, P. D. (2007). An Instinct for Detection: Psychological Perspectives on CCTV Surveillance. *The Police Journal*, 80, 287-305.

- Scott-Brown, K. C., & Mann, D. P. (2005). On the costs and benefits of display format in a video-based observation task. *SProceedings of the 2<sup>nd</sup> Symposium on Applied Perception in Graphics and Visualization*, 95 Archive, 157.
- See, J. E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin*, 117, 230-240.
- Shen, Y. J., & Jiang, Y. V. (2006). Interrupted visual searches reveal volatile search memory. *Journal of Experimental Psychology: Human Perception & Performance*, 32, 1208-1220.
- Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist*, 61(4), 394-403.
- 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. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, 5, 644-649.
- Simmonds, D. C., Poulton, E. C., & Tickner, A. H. (1975). Identifying people in a videotape recording made at night. *Ergonomics*, 18(6), 607-618.
- Snyder, H. L. (1974). Image quality and face recognition on a television display. *Human Factors*, 16, 300-307.
- Snyder, M., & Swann, W. B. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, 36, 1202-1212.
- Sohn, Y. W., Doane, S. M., & Garrison, T. (2006). The impact of individual differences and learning context on strategic skill acquisition and transfer. *Learning and Individual Differences*, 16, 13-30.
- Speier, C., Valacich, J. S., & Vessey, I. (1999). The influence of task interruption on individual decision making: An information overload perspective. *Decision Sciences*, 30, 337-360.
- Sperling, G. (1981) Video transmission of American Sign Language (ASL) and Fingerspelling: Present and projected bandwidth requirements. *IEEE Transactions on Communications*, com-29(12), 1993-2002.
- Stebly, N. J. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior*, 16, 413-424.
- Stone, J. V. (1999). Object recognition: View-specificity and motion-specificity. *Vision Research*, 39, 4032-4044.
- Stone, J. V., Buckley, D., & Moger, F. A. (2000). Determinants of object recognition. *Vision Research*, 40, 2723-2736.
- Swann, W. B., Jr., Giuliano, T., & Wegner, D. M. (1982). Where leading questions can lead: The power of conjecture in social interaction. *Journal of Personality and Social Psychology*, 42, 1025-1035.

- Tanaka, J. W., & Farah, M. J. (1993) Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, 46, 225-245.
- Tarr, M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience*, 1, 275-277.
- Tesser, A., Martin, L., & Mendolia, M. (1995). The impact of thought on attitude extremity and attitude-behavior consistency. In R. E. Petty & J. A. Krosnick (Eds.), *Attitudes strength: Antecedents and consequences* (pp. 73–92). Mahwah, NJ: Erlbaum.
- Thomas, M. (1993) *Every Mother's Nightmare: The Killing of James Bulger*. London: Pan.
- Tickner, A. H., & Poulton, E. C. (1968). Remote monitoring of motorways using closed-circuit television. *Ergonomics*, 11(5), 455-466.
- Tickner, A. H., & Poulton, E. C. (1973). Monitoring up to 16 synthetic television pictures showing a great deal of movement. *Ergonomics*, 16(4), 381-401.
- Tickner, A. H., & Poulton, E. C. (1975). Watching for people and actions. *Ergonomics*, 18(1), 35-51.
- Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies*, 58, 583–603.
- Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., & Wilson, A. (2004). What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception*, 33, 87-101.
- Tuchman, B. W. (1984). *The march of folly: From Troy to Vietnam*. New York: Ballantine Books.
- Tversky, A., & Kahneman, D. (2002). Judgment under uncertainty: Heuristics and biases. In D. J. Levetin (Ed.). *Foundations of cognitive psychology: Core readings*. (pp. 585-600). Cambridge, MA, US: MIT Press.
- U. S. Department of Labor (2008). Workstation Environment.  
[http://www.osha.gov/SLTC/etools/computerworkstations/wkstation\\_enviro.html#glare](http://www.osha.gov/SLTC/etools/computerworkstations/wkstation_enviro.html#glare) .
- Vasterling, J. J., Brailey, K., Constans, J. I., & Sutker, P. B. (1998). Attention and memory dysfunction in posttraumatic stress disorder. *Neuropsychology*, 12(1), 125-33.
- Veiel, L. L., Storandt, M., & Abrams, R. A. (2006). Visual search for change in older adults. *Psychology & Aging*, 21, 754-762.
- Von Bueren, J. A., Radanov, B. P., & Jäncke, L. (2005). Pain Influences Several Levels of Attention. *Zeitschrift für Neuropsychologie*, 16(4), 235-242.
- Voss, J. F., & Post, T. A. (1988). On the solving of ill-structured problems. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.). *The Nature of Expertise*. (pp. 261-285). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

- Warm, J. S., & Dember, W. N. (1998). Tests of vigilance taxonomy. In R. R. Hoffman, M. F. Sherrick, & J. S. Warm (Eds.) *Viewing Psychology as a Whole: The Integrative Science of William N. Dember*. (pp. 87-112). Washington, DC, US: 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* (pp. 183-200). Mahwah, NJ: Erlbaum.
- Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.
- Washburn, D. A., Tagliatela, L. A., Rice, P. R., & Smith, J. D. (2004). Individual Differences in Sustained Attention and Threat Detection. *International Journal of Cognitive Technology*, 9(2), 30-33.
- Weick, K. (1999). Sensemaking as an organizational dimension of global change. In D. Cooperrider, & J. Dutton (Eds.), *Organizational Dimensions of Global Change: No Limits to Cooperation* (pp. 39-56). CA: Sage Publications.
- Weinger, M. B., & Smith, N. T. (1997). Vigilance, Alarms, and Integrated Monitoring Systems. Unpublished paper retrieved from <http://www.mc.vanderbilt.edu/cprq/publications.html>.
- Wells, G. L. & Olson, E. (2003). Eyewitness identification. *Annual Review of Psychology*, 54, 277-295.
- Werner, S., & Thies, B. (2000). Is change blindness attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition*, 7, 163-173.
- Wickens, C. D., & Dixon, S. R. (2007). The benefits of imperfect diagnostic automation: A synthesis of the literature. *Theoretical Issues in Ergonomics Science*, 8(3), 201-212.
- Wijesekera, D., Srivastava, J., Nerode, A., & Foresti, M. (1999). Experimental evaluation of loss perception in continuous media. *Multimedia Systems*, 7, 486-499.
- Wilson, G., & Russell, C. (2003). Operator functional state classification using multiple psychophysiological features in an air traffic control task. *Human Factors*, 45, 381-389.
- Witkin, H. A. (1964). Origins of cognitive style. In C. Sheerer (Ed.), *Cognition: Theory, Research and Promise* (pp. 172-205). New York: Harper & Row.
- Woelders, W. W., Frowein, H. W., Nielsen, F., Questa, P., & Sandini, G. (1997). New developments in low-bit rate videotelephony for people who are deaf. *Journal of Speech & Hearing Research*, 40(6), 1425-1433.
- Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, 1(2), 202-238.

- Wright, D. B., & McDaid, A. T. (1996). Comparing system and estimator variables using data from real line-ups. *Applied Cognitive Psychology*, 10, 75-84.
- Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica*, 114, 101-114.
- Yamazaki, T. (2001). Subjective video assessment for adaptive quality of service control. *Proceedings of the IEEE International Conference on Multimedia and Expo*. Tokyo, Japan. pp. 517-520.
- Yoo, S. K., Kim, D. K., Jung, S. M., Kim, E.-K., Lim, J. S., & Kim, J. H. (2004). Performance of a Web-based, realtime, tele-ultrasound consultation system over high-speed commercial telecommunication lines. *Journal of Telemedicine and Telecare*, 10, 175-179.
- Zacks, J., Tversky, B., & Iyer, G. (2001). Perceiving, remembering and communicating structure in events. *The Journal of Experimental Psychology: General*, 130, 29-58.
- Zijlstra, F. R. H., Roe, R. A., Leonora, A. B., & Krediet, I. (1999). Temporal factors in mental work: Effects of interrupted activities. *Journal of Occupational and Organizational Psychology*, 72, 163-185.
- Zsombok, C. E. & Klein, G. (Eds). (1997). *Naturalistic decision making*. Mahwah, NJ: Lawrence Erlbaum Associates.