



UNIVERSITY OF
LINCOLN

**Paying Attention to the Evidence: A Comparison of Perception
and Decision Making Processes in Novice and Experienced
Scene of Crime Officers Using Eye Tracking in Simulated Crime
Scene Scenarios**

Murat Ozger

A thesis submitted in partial fulfilment of the requirements of the University of Lincoln for
the degree of Doctor of Philosophy

July 2016

Abstract

Research on crime scene investigation has strongly focused on the technical aspects of the process, while cognitive aspects (searching, reasoning and perception) have often been overlooked. Textbooks on forensic sciences tend to focus on identifying and processing evidence, and the use of equipment while it can be argued that cognitive factors in processing such evidence and using equipment are equally important. This thesis studies the cognitive aspects of crime scene investigation by comparing eye movement patterns in experts and novices. Studies in various domains, including surgery, sports, and chess playing have shown that eye movements differ between experts and novices, providing a tool towards a more objective assessment of skill than is possible with peer assessment. In four experiments eye movements of experts and novices were examined during (1) inspection of photographs of crime scenes on a computer screen (2) a change blindness task on crime and non-crime scene images, (3) active exploration of a simulated crime scene and (4) the assessment of emotional crime and natural scenes. While some trends in eye movement differences, such as a tendency on longer fixation durations and a broader focus on the overall scene and less on the direct evidence could be found in experts compared to novices, differences between experts and novices were considerably smaller than in other domains, despite the broad range of measures extracted from the data. This lack of clear expertise effects may relate to the rather diverse range of perceptual layouts of crime scenes, reducing possible top-down effects of expertise on the deployment of attention. The results will be discussed with a view of possible directions of future research in this domain.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisors Prof. Timothy Hodgson, Dr. Jose Gonzalez-Rodriguez and Dr. Frouke Hermens for the continuous support of my PhD for their patience, motivation and immense knowledge. I could not have imagined having a better advisor and mentor group for my PhD study.

My deepest gratitude also goes to all my friends especially in Lincoln and of course in Turkey for their priceless support and also encouragement to finish my PhD study.

Special thanks go to my colleague Jeremy Smith who let me conduct the studies with crime scene investigators in his department despite their heavy work load. I am grateful to have collaborated with a different police department and get to know all the great people with whom i hope to keep in touch in the future.

Finally, I want to thank to my beautiful wife Hande and my Mother, Father and Sister, for believing in me, being with me and supporting me all these three years although I was far away from them. I love you all so much.

Table of Contents

Abstract	ii
Acknowledgments	iii
Table of Contents	iv
List of Figures	x
List of Tables	xiii
List of Abbreviations	xiv
Chapter 1. Crime Scene Examination	1
1.1. Introduction	1
1.2. Crime Scene Examination	3
Chapter 2. Information Processing	15
2.1. Introduction	15
2.2. Information Processing Approach.....	15
2.3. Fallibility in Perception	22
2.4. Emotion as a Factor in Information Processing	24
2.5. Experts and Novice Paradigm	27
2.6. Visual Expertise.....	31
2.6.1. Sport.....	32
2.6.2. Chess	33
2.6.3. Aviation	34
2.6.4. Problem Solving	35
2.6.5. Medicine	36
2.6.6. Driving	37

2.6.7. Forensic Sciences.....	38
2.6.7.1. Studies on Forensic Lab Examinations	38
2.6.7.2. Eye-Witness Work as a Part of Criminal Investigations	42
2.7. Conclusions	43
Chapter 3. Methods.....	46
3.1. Introduction	46
3.2. History of Eye-Tracking.....	46
3.3. Types of Movements	49
3.3.1. Fixations and Saccades	49
3.3.2. Fixational Eye Movements	50
3.3.3. Smooth Pursuit and Opto-Kinetic Nystagmus.....	51
3.3.4. Vestibular Oculomotor Reflex.....	51
3.3.5. Vergence Eye Movements	52
3.4. Most Common Eye Movement Measures	52
3.4.1. Dwell Time	54
3.4.2. Number of Fixations	54
3.4.3. Average Fixation Duration	56
3.4.4. Time to First Fixation	56
3.4.5. Saccade Amplitude	57
3.4.6. Saccadic Latencies.....	57
3.4.7. Scanpath Measures	58
3.4.8. Blinks	58
3.4.9. Pupil Dilation.....	58
3.4.10. Heatmaps	59
3.4.11. Recurrence Quantification Analysis	60

3.5. Eye Tracking Techniques	62
3.5.1. Electro-OculoGraphy.....	62
3.5.2. Scleral Coil	63
3.5.3. Dual Purkinje Image Eye Trackers	64
3.5.4. Limbus Eye Trackers.....	65
3.5.5. Dark and Light Pupil Methods	66
3.5.6. Mobile Eye Trackers.....	67
3.5.7. Combining Methods	68
3.6. Equipment in current study	69
3.6.1. The Eyelink 1000 Desk Mounted Eye Tracking Device	69
3.6.2. The Positive Science Mobile Eye Tracking Device	72
3.7. Non-eye movement methods	74
3.8. Conclusion.....	77
Chapter 4. Eye Movements in Experts and Novice Crime Scene Investigators during	
Viewing of Images of Simulated Crime Scenes	78
4.1. Introduction	78
4.2. Method.....	81
4.2.1. Participants	81
4.2.2. Ethics Statements	82
4.2.3. Materials	82
4.2.4. Apparatus	83
4.2.5. Procedure and Design	86
4.2.6. Data Analysis.....	88
4.2.7. Results.....	93
4.2.7.1. Eye Movement Characteristics	93

4.2.7.2. Selected Evidence	101
4.2.7.3. Verbal Reports	102
4.2.7.5. Case Hypothesis – Role of scenario.....	106
4.3. Discussion.....	106
4.4. Conclusion.....	113
Chapter 5. Comparing expert and non-expert CSIs in a Change Blindness	
Experiment.....	114
5.1. Introduction	114
5.2. Method.....	118
5.2.1. Participants	118
5.2.2. Ethics Statements	119
5.2.3. Materials	119
5.2.4. Apparatus.....	120
5.2.5. Procedure and Design	120
5.2.6. Data Analysis.....	123
5.2.7. Results.....	124
5.2.7.1. Performance Data.....	124
5.2.7.2. Eye Movement Characteristics	125
5.3. Discussion.....	127
5.4. Conclusion.....	131
Chapter 6. Eye Movements in Volume and Senior Crime Scene Investigators during	
Active Exploration of a Simulated Crime Scene	133
6.1. Introduction	133
6.2. Method.....	137
6.2.1. Participants	137

6.2.2. Ethics Statements	137
6.2.3. Materials	138
6.2.4. Apparatus	140
6.2.5. Procedure and Design	140
6.2.6. Data Analysis	142
6.2.7. Data Quality	147
6.2.8. Results	147
6.2.8.1. Eye Movement Characteristics	148
6.2.8.2. Verbal Reports	158
6.2.8.2. Prioritized Evidence	161
6.2.8.5. Case Hypothesis – Role of scenario	162
6.3. Discussion	163
6.4. Conclusion	170
Chapter 7. Effects of Training on Behavioral and Evaluative Responses during	
Viewing Emotional Stimuli	171
7.1. Introduction	171
7.2. Method	176
7.2.1. Participants	176
7.2.2. Ethics Statements	176
7.2.3. Materials	176
7.2.4. Apparatus	178
7.2.5. Procedure and Design	178
7.2.6. Data Analysis	180
7.2.7. Results	182
7.2.7.1. Pupil Responses	182

7.2.7.2. Eye Movement Characteristics	186
7.2.7.3. Memory Performance	187
7.2.7.4. Valence and Arousal Ratings	189
7.3. Discussion.....	192
7.4. Conclusion.....	199
Chapter 8. Conclusions	200
8.1. Background.....	200
8.2. Summary of the Findings	202
8.3. Implications of the Findings	205
8.4. Strengths and Weaknesses of the Experiments	209
8.5. Recommendations for Further Research	213
8.5.1. Focusing on Team Work	213
8.5.2. Real Case Work	214
8.5.3. Use Eye Tracking for Feedback and Self-Reflection	214
8.5.4. Additional Tools to Explore The Processes.....	216
8.6. Conclusion.....	217
List of References	219
Appendix-A. Invitation Letter to Managers	284
Appendix-B. List of Details of the Stimuli in Chapter 5	286
Appendix-C. List of Details of the Stimuli in Chapter 7	287

List of Figures

<i>Figure 1.1.</i> The number of yearly citations in contextual and confirmation bias in all sciences and forensic sciences per year.....	6
<i>Figure 1.2.</i> Illustration of the crime scene investigation approach.....	11
<i>Figure 2.1.</i> Illustration of an information processing model	16
<i>Figure 3.1.</i> An example of an heatmap from the study.....	60
<i>Figure 3.2.</i> An example of an earlier eye movement recording device and recordings	63
<i>Figure 3.3.</i> An image of an Eyelink 1000 desk mounted eye tracking system.....	70
<i>Figure 3.4.</i> Illustration of the Positive Science mobile eye tracker system	73
<i>Figure 3.5.</i> Illustration of the Yarbus software.....	74
<i>Figure 4.1.</i> Images used for the three scenarios in the study.....	84
<i>Figure 4.2.</i> AOIs used for analysis for the three scenarios	89
<i>Figure 4.3.</i> Illustration of the fixed grid method.....	91
<i>Figure 4.4.</i> Illustration of the fixation distance method.....	91
<i>Figure 4.5.</i> Plots of average eye movement data for each image and scenario and participant group	96
<i>Figure 4.6.</i> Example of fixation sequences and recurrence plot using fixation distance method.....	98
<i>Figure 4.7.</i> Plots of average CRQA measures for each image between and within group	100
<i>Figure 4.8.</i> Number of objects selected as evidence for each scenario and participant group.....	101
<i>Figure 4.9.</i> Overall proportions of coded retrospective reports in each participant group.....	103

<i>Figure 4.10.</i> Plots of average references stated in verbal reports for each scenario between groups	105
<i>Figure 5.1.</i> The stimulus sequence	122
<i>Figure 5.2.</i> Example of a simulated crime scene change	123
<i>Figure 5.3.</i> Example of a neutral scene change	123
<i>Figure 5.4.</i> Plots of average accuracy of responses, response times, fixation durations, numbers of fixations and dwell times on accurate change detection trials for participants groups	126
<i>Figure 6.1.</i> Photographs illustrating the simulated crime scene	138
<i>Figure 6.2.</i> The layout of the Coded AOIs	144
<i>Figure 6.3.</i> Snapshot of a participant’s point of view in the study	148
<i>Figure 6.4.</i> Plots of average cumulative dwell times on placed AOIs and entry and exit points	149
<i>Figure 6.5.</i> Plots of average cumulative dwell times on uncategorized items and possible evidence AOIs	151
<i>Figure 6.6.</i> Scatter plots showing the association between years of experience and dwell times on AOIs	153
<i>Figure 6.7.</i> Plots of between group and within group average CRQA measures	157
<i>Figure 6.8.</i> Plot of average total durations between scenarios and experience.	158
<i>Figure 6.9.</i> Plot of stated references in the verbal reports	160
<i>Figure 6.10.</i> Prioritized evidence across participant groups and scenarios	162
<i>Figure 6.11.</i> Generated hypothesis for participant groups on each scenario	163
<i>Figure 7.1.</i> Example of negative stimuli	181
<i>Figure 7.2.</i> Example of neutral stimuli	182

Figure 7.3. Plots of average peak dilations(au) and peak dilaion latency for each emotional content on image types per participant groups.....183

Figure 7.4. Plots of average peak dilations (au) and peak dilation latency for the initial 2 seconds for each emotional content on image types per participant groups185

Figure 7.5. Plots of average number of fixations and saccade amplitudes for each emotional content on image types per participant groups.....187

Figure 7.6. Plots of average memory accuracy and Confidence Ratings for each emotional content on image types per participant groups.....189

Figure 7.7. Plots of average valence, arousal and complexity ratings for each emotional content on image types per participant groups.....191

Figure 7.8. Scatter plot comparing peak dilations and subjective ratings for a range of ratings (valence, arousal and complexity).192

Figure 8.1. The information processing and reasoning process in crime scene examination process.....208

List of Tables

<i>Table 2.1.</i> The underlying perceptual and cognitive processes during the crime scene examination process	22
<i>Table 2.2.</i> Studies that used eye tracking technology in traditional forensic disciplines comparing the domain of research and the eye tracking measures used.....	40
<i>Table 4.1.</i> The three scenarios which were presented to the participants before the two images for each case.....	85
<i>Table 4.2.</i> An example of coding process	104
<i>Table 6.1.</i> The two scenarios which were presented to the participants before entering the scene	139
<i>Table 6.2.</i> Coded AOIs in the Scene	145
<i>Table 6.3.</i> An example of coding process	159

List of Abbreviations

Abbreviation	Meaning
AOI	Area of Interest
AU	Arbitrary Unit
CRQA	Cross Recurrence Quantification Analysis
CSI	Crime Scene Investigator
DIEM	Dynamic Images and Eye Movements
DPI	Dual Purkinje Image
EOG	Electro-OculoGraphy
FBI	Federal Bureau of Investigation
IREM	Infrared Emitting Diode
LTM	Long Term Memory
M	Mean
ms	Millisecond
NAS	National Academy of Sciences
OKN	Opto-Kinetic Nystagmus
PTSD	Post-traumatic stress disorder
RQA	Recurrence Quantification Analysis
s	Second
SD	Standard Deviation
SE	Standard Error
STM	Short Term Memory
USA	United States of America
VOR	Vestibular Oculomotor Reflex

Chapter 1. Crime Scene Examination

1.1. Introduction

In his book about homicide investigations, Innes (2003) explained criminal investigation as: “The identification, interpretation and ordering of information with the objective of ascertaining whether a crime has occurred, and if so, who was involved and how” (p. 113). This quote illustrates the challenges and complexities associated with crime scene investigation. Mol and Mesman further illustrate these complexities as crime scene “site[s] where people belonging to different worlds and talking different languages gather” (as cited in Ludwig, Fraser & Williams, 2012, p.58). In this quote they indicate how all players contributing to investigation processes must work together, even while each having their own responsibility. The ultimate aim of the crime scene investigation process is to shed light on the crimes and contribute to the mission of accurate and fair justice outcomes (Julian, Kelty & Robertson, 2012; Kelty, Julian & Robertson, 2011).

For decades, the crime scene investigation process has been based on experience in past cases, but recently, more and more influences of scientific practice have been brought into crime scene investigation. This has been caused by the introduction of evidence-based approaches, bringing major changes into policing (Sherman, 2013). As a result, reactive policing strategies in the late 70s shifted to proactive strategies to maintain public safety (Sherman, 2013).

With changing attitudes towards applying the scientific approach to policing, forensic science has become much more crucial and influential over the last few decades. The advancement of DNA technologies and computing technology allowing the search of more and more extensive databases have also contributed to the new policing strategies. Television shows on crime scene investigation have further affected the overall attitude of the public and generated great interest on forensic sciences (Wilson-Kovacs, 2013).

Crime scene investigation is considered to be the initial phase of the process that provides information or facts about perpetrators or incident course of events in a scientific manner (Julian et al., 2012). The whole process starting from the crime scene is known as the forensic investigation (Schiffer, 2009). This process is based on the idea that "Every contact leaves a trace", known as the Locard's principle (Horswell, 2004), although how it got linked to the French scientist Edmond Locard is unclear, as it is claimed he has never made such statement (Wyatt, 2014). Evidence collected from crime scenes not only helps to understand the activities related to a case but also helps to link crime scenes or crimes. Linking crime scenes with one another, identifying suspects and gathering more information on how crimes are committed are part of proactive policing strategies, which is also known as intelligence led policing (Ribaux et al., 2010a; Ribaux et al., 2010b).

In order to conduct an effective investigation, domain-specific knowledge, experience, skills and attitudes play a major role in addition to the technologies that aid the investigation (Kelty et al., 2011). A full grasp of the cognitive processes involved, including perception, reasoning, decision making and training, experience and coordinated team work, is essential for quality crime scene investigation (Lee & Pagliaro, 2013). In different steps of the crime scene examination process, cognitive processes play an important role, in addition to motor

skills (such as collecting and packaging traces at the crime scene; Lee & Pagliaro, 2013), although they have rarely been addressed by empirical research. In the majority of books and articles on crime scene investigation and forensic science, attitudes and cognitive process of competent crime scene investigators have not been discussed until recently (Dror, 2015; Kelty et al., 2011; Lee & Pagliaro, 2013). Due to the complex nature of crime scene investigation, technological support has always taken priority and very few studies have been conducted on crime scene investigators' perception and cognitive processes, skills, attitudes or training. In this section, I will try to concentrate on the importance of these processes and will also explore the few studies which have been done in the domain.

1.2. Crime Scene Examination

The bigger picture during an investigation consists of three phases: (1) the crime scene examination (finding and collecting traces), (2) laboratory analyzes (analysis of the traces) and (3) writing reports (evaluation and interpretation of the results) (Baber, 2010; Schiffer, 2009). These processes are considered as part of the investigative and evaluative work within the criminal investigation (Jackson, Aitken & Roberts, 2015). Generally, the majority of textbooks and training courses focus on the identification and recovery of important traces (evidence) in a case, the documentation and analysis of forensic evidence (e.g. Fisher & Fisher, 2012; Horswell, 2004). The evidence which is often referred to as forensic artefacts, include DNA profiles, fingerprints, shoe marks, blood patterns, and tool marks (Williams & Weetman, 2013). In other documents, forensic artefacts are referred to as traces (Horswell, 2004; Wyatt, 2014). A related focus of classical texts is on the collaboration with other people entering the scene, such as police officers and paramedics. Such collaboration is essential as

others in the scene can provide relevant information but also disturb the site (Kelty et al., 2011).

For the last decade, there has been an increasing number of erroneous judicial cases from the past and there is a huge criticism against forensic sciences and crime scene examinations. Undoubtedly, insufficiently examined crime scene/s lead to several problems about the investigations of the cases (Dror, 2015; Lee & Pagliaro, 2013). Any professionally made error, failure and/or negligence of recognizing and recovering evidence along with the contamination issues at the crime scene or the laboratory environment in the forensic investigation make the acquired analysis and results ineffective and invalid (Lee & Pagliaro, 2013). As a consequence, contaminated or poorly conducted forensic processes starting from the crime scene shall lead to wrongful outcomes, such as the failure to convict a perpetrator or the conviction of wrong person, sometimes referred to as “rubbish in - rubbish out” (Ribaux et al., 2010a, p.11).

In order to display an outstanding performance, there are several issues such as human factors that can withhold or hinder the desired outcome. In most forensic domains from crime scene examination to analysis of evidence in forensic laboratories, human performance is the core element of the work. The expert who perceives information, interprets its relevance and makes judgments about it, is the “instrument of analysis” (Dror, 2015, p.40). However, the human brain has limited processing capacity. For this reason, not all information that enters through our senses is encoded and processed, which has implications for human performance. Combining the domain knowledge with the procedural guidelines shall help the experts to finish the examination process in an effective and efficient way, but also deliberate practices of different cases shall lead the examiners to improve their performances at the crime scenes

(Ericsson, 2006; Ericsson, Krampe & Tesch-Romer, 1993; Ericsson & Towne, 2010). Due to this, experience is a very important at the crime scene and mostly, experienced examiners are sent to the scenes to conduct crime scene examinations (Kelty et al., 2011). The experience of the examiner and his skills will lead to high quality assessment, interpretation and processing at the crime scene.

Despite its importance, the required skills, attitudes and cognitive process of competent crime scene investigator or forensic scientist are almost neglected and not discussed in the literature (Julian et al., 2012; Kelty, 2011). The publication of the National Academy of Sciences (NAS) Report (2009), which can also be described as the latest criticism of the lack of principles and guidelines in forensic sciences, led to an increased number of studies focusing on cognitive processes. After misidentification of the fingerprints of the suspect in the Madrid bombing case (2004) by the Federal Bureau of Investigation (FBI) in United States of America (USA), significant numbers of studies have been conducted on how to best evaluate forensic evidence by the experts focusing by means of human factors but mainly on explaining the effects of contextual information and bias in expert's decisions (Found, 2015; Kassin, Dror & Kukocka, 2013 for review; Nakhaeizadeh, Dror & Morgan, 2014). Studies such as these found that providing contextual information to experts biased their investigation processes, while others have suggested no effect of contextual information (Kerstholt, Paashuis & Sjerps, 2007). These investigations have provided vital insights about the reasons of the errors, the factors that define expertise and the importance of multidisciplinary research. The studies also suggest that by concentrating merely on the outcomes of expert performance has destructive effects on forensic sciences (Champod, 2014).

Despite a recent increase in the number of studies examining cognitive processes involved in forensic sciences, the overall number of studies on this aspect is still not promising (Found, 2015). In his study, Found (2015) focused his comparison of forensic sciences and other disciplines on the effect of contextual information and confirmation bias (see Figure 1.1). These two concepts have only been introduced to forensic textbooks or scientific journals recently. For instance, in the recent book “Professional Issues in Forensic Science” (2015), Schuliar and Crispino provided an introduction into the reasoning and decision making processes for forensic scientists. Further details were described by Smith and Bond (2014) focusing on inferential reasoning in the real cases. Earlier work by Inman and Rudin (2000) highlighted the need to study cognitive factors. And recent work by Dror (2015) emphasized the neglected nature of human factors in forensic sciences and the need to conduct more empirical research on the perception and cognition in crime scene investigation.

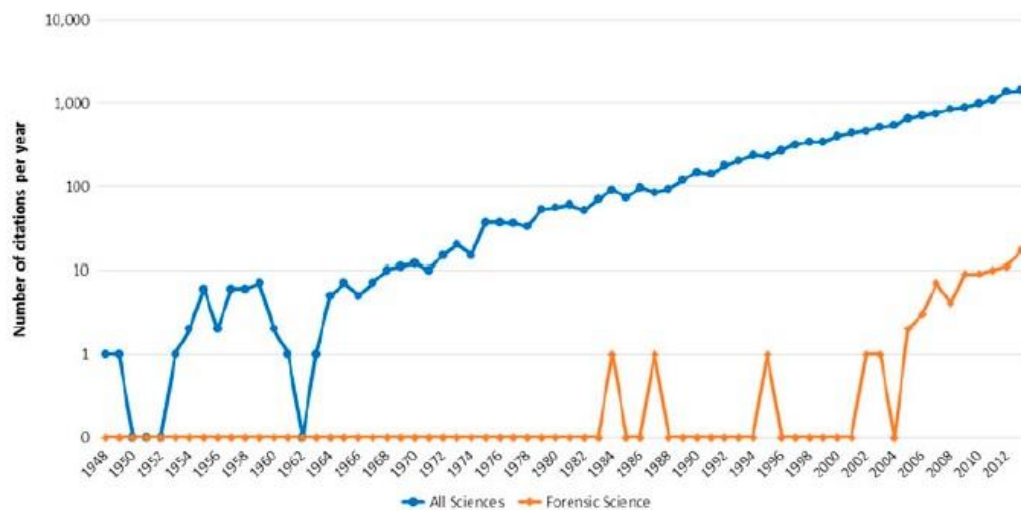


Figure 1.1. The number of yearly citations of contextual and confirmation bias in all sciences (blue line) and forensic sciences (orange line) (log scale, adapted from Found, 2015, p.396).

Studies on what defines expertise in crime scene investigation are sparse, in particular when concerned with cognitive aspects. Besides, lack of professional and training standards for crime scene investigators still remain to be one of the challenges encountered in this field. In different corners of the world, different working groups have been founded to advise forensic sciences and CSIs (see Houck, 2015). These working groups aim to adapt new procedures to crime scene investigations and other forensic disciplines. Their core activities also include the standardization of the techniques applied at the crime scene and laboratories, and drafting and publishing practice manuals. Despite these aims standardization studies which scrutinize the errors made at crime scenes are still lacking. Moreover, methods to improve performance of CSIs in the crime scenes remain unexplored (Kelty et al., 2011). Furthermore, studies are still needed that examine skill acquisition, evaluation of CSIs, and the roles of structured and standardized training. While past scientific studies have undeniably improved the quality of crime scene investigation, new horizons which integrate cognitive studies bring more efficiency and effectiveness.

The fact that every crime scene is different may explain why there have been fewer studies of the actual crime scene investigation process. While limited in number, studies of crime scene investigation have examined a broad range of research methods. Earlier studies have focused predominantly on the implementation of new methods or new technologies. Recent work, in contrast have started to examine what determines expertise in crime scene investigation.

Schraagen and Leijenhorst (2001) were one of the first to compare performance of experts and novices, although the experiment focused merely on the recovery of evidence from clothing. In their study, they studied the process of collecting evidence in a forensic

laboratory setting. Experience was used as a proxy for expertise. Experts were defined with four and six years of experience in laboratory settings, while the novice participant was the intern working in the laboratory for 6 months or less. The results of their study suggested that time available for evidence collection and adequate advance information about the case were crucial in the search process. The study shed light on the individual search strategies in exploration of evidence among the participants. The study did support the view that increasing experience helps the search for evidence and the construction of hypotheses about the case.

A recent study observing and comparing expert and novice performance highlighted the importance of experience and training for search and recovery strategies in crime scenes (Baber & Butler, 2012). In their experiments, Baber and Butler made use of a “point of view” camera that records a video file from the point of view of the person wearing the camera (c.f., Omodei & Mc Lennan, 1995), without tracking the viewer’s eye movements. The use of this camera demonstrated differences in search strategies between undergraduate forensic students and experienced CSIs in two simulated burglary scenarios. Experts finished the task more quickly and formed better narratives about the crime, suggesting that experts better judge the relevance of the items in the crime scene and their evidential value for future analysis in the laboratory. Experienced CSIs also appeared to use a more holistic approach in developing further plans in the investigation process. In contrast, students concentrated more on the context of the simulated scenarios. One could argue that expertise in crime scene investigation should not be measured as the number of years in the job, but on the other hand, it is plausible that more experienced CSIs also demonstrate more expertise (e.g., Keltz et al., 2011). While experts generally performed better than novices, they still were found to make errors (Baber & Butler, 2012).

Another dimension on which expert CSIs may differ from novices is decision making. Evidence for this was found by Helsloot and Groenendaal (2011), who examined decision making processes of Dutch forensic team leaders. Using a focus group and different case studies, they explored decision making mechanisms in crime scenes. They scrutinized the interaction with contextual information and posited that in crime scene work, team leaders mostly based their decisions on background information, particularly in emotional cases (e.g. murder of a child). This demonstrated the potential for bias in experts' decision making and judgment processes. Besides background information, available information, personal traits and time are important factors in the decision making process in forensic team leaders.

Factors, such as task complexity, type of crime, context, time and the environment, influence the cognitive aspects of the crime scene investigation process (Morrison, Wiggins & Porter, 2010). Experts are thought of being better at dealing with these various aspects of the work. While evaluating the design of a support system to improve CSIs cognitive skills in perception, interpretation and decision making, Morrison et al. (2010) found that expert CSIs tended to acquire information more easily than novices.

Further evidence towards the nature of differences between experts and novices was found by Wilson-Kovacs (2013) who interviewed "experienced" crime scene personnel, exploring their understanding of the investigative practice and profession. In her interviews, the expert interviewees' indicated to more easily spot the objects out of place. They also reported to be able to rely more strongly on experience than their novice counterparts.

Additional characteristics of expertise in CSI are enhanced knowledge, cognitive skills, work orientation, life experience, communication skills, professional manner and approach (Kelty et al., 2011). Their data suggest that years of experience alone is not sufficient for high performance, and will not guarantee the attainment of expert performance. On the basis of the skills identified, they constructed recommendations for the recruitment of personnel (Kelty, 2011). Such recruitment strategies could involve specific tests and interviews focused in the specific skills identified in experts (Kelty & Gordon, 2012). It is yet unclear whether such recruitment strategies extend to other domains in forensic sciences, although it has been suggested that in the recruitment of border personnel it might be beneficial to screen for “super-recognizers”, with a natural talent for recognizing faces (Robertson, Noyes, Dowsett, Jenkins & Burton, 2016).

The crime scene examination process engages with all objects at the scene and the information about the case using systematic methods to develop explanations or hypotheses to find out the answers to "who done it?" and "how done it?" questions (Baber, 2010; Jackson, et al., 2015). Two manuals, part of the recently published Practitioner Guide, Roberts and Aitken (2014) and Jackson et al. (2015) attempt to formulate a framework for case assessment in forensic science, crime scene investigation, law cases using a statistical approach. These manuals focus on the “probative value” of the evidence (Lee & Pagliaro, 2013, p.2). Roberts and Aitken (2014) compared the cognitive reasoning process to a specialized inferential reasoning process, applying a scientific approach to using evidence (information) to reach conclusions about the case. They divide the case assessment process into two aspects as afore stated above: (1) Investigative work and (2) evaluative work, the latter also referred to as lab work. The basic task of the CSIs is to reduce the ambiguity of the cases by applying the inferential reasoning forms such as deductive, inductive and abductive reasoning on the basis

of the existing evidence during the investigative work (Jackson et al., 2015). Inductive reasoning can be defined as a bottom up process from observations to theory and deductive reasoning as the top down process, theory leading the observations to reach the conclusions (Baber, 2010). Experience-based and domain knowledge will particularly help the abductive reasoning process, in which information is evaluated and in which an initial hypothesis is formed (Jackson et al. 2015). The abductive reasoning process is a mixed analytical and assumption process, normally involving formulating explanations on the basis of the information and known probabilities. The outcomes of this dynamic process can be altered by new information, facts and deductive inferences.

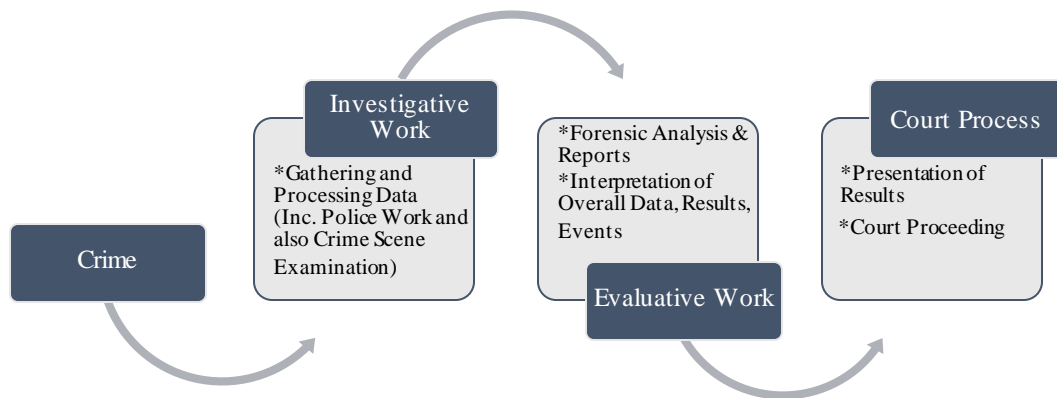


Figure 1.2. Illustration of the crime investigation approach. After a crime, investigative work is conducted, involving the gathering of data, which is then evaluated and presented in court.

In the past, doubts have been cast on the importance of improving the crime scene investigation process (Kelty, 2011). One of the criticism was that the process was seen only as a "bagging and tagging" exercise, mainly based on inductive inferences (Harrison, 2006; Julian et al., 2012). Such a view is surprising given the portrayal of detective work in the popular literature, for example in Conan Doyle's novels, where the Sherlock Holmes

character's ability to create different hypotheses from his observations in the crime scene and reach conclusions from eliminating, all depended on having sufficient information about the case (Carson, 2009; Crispino, 2008).

Reasoning in crime scene work depends heavily on artefacts, the environment and domain knowledge. The use of domain knowledge in accordance with the procedural guidelines constitutes a fundamental part of the process. As all cases are different and have unique characteristics, it is at the discretion of the CSI to focus on the relevant items and specify and gather evidence. Collecting any artefact at the crime scene regardless of the forensic value leads to an increase in the workload of crime scene examiners and consequently will have a negative effect on the lab analysis. When acting at a crime scene, the examiners should bear in mind the evidential value or "probative value" (Lee & Pagliaro, 2013). Evidence collected at the scene of crime results in elimination of diverse hypotheses and possible scenarios leading to more concrete outcomes. In this respect, choosing evidence that has probative value facilitates and shapes the course of investigation. Therefore, probative value can be of a great benefit while generating hypothesis or linking the scene to a specific suspect (Lee & Pagliaro, 2013; Schiffer, 2009).

In his study, Baber (2010) indicated the inseparable link of crime scene work from the forensic sciences and explored how crime scene examiners' cognition and performance interacts with the complexity of the case. In his study, he focused on the different stages of the forensic process and the roles of the different actors in the process and the information flow between them. In his study, he did not only focus on the interaction between the scene and the examiner, but also explained how the other actors have influence on the mental states of the examiners. He revealed the different stages of the crime scene examination process

(e.g., recovery, collection, documenting) and the relationship between different actors (e.g., first responders, investigators, forensic scientists). He defined all these influences as the resources for actions and defined the cognitive processes of the examiner in terms of these resources. For example, perception of the relevance of an artefact in the scene will lead to consider different responses and actions, and interpretations will largely depend on existing knowledge and experience. The different roles of the actors in the whole process and the interactions between these actors will influence the quality of the outcomes. He concluded that the crime scene process should look beyond the processing and documenting of evidence, and should instead be examined in the context of the actors in the process.

Further research on how crime scene investigators approach a (simulated crime scene) has focused on the use of new technologies (De Gruijter et al., 2016). In this study, CSIs were presented with new technologies, such as a spectral camera to detect invisible blood traces, which were often new to the crime scene investigator who was asked to use these new devices. Other new technology tested included a mobile DNA analysis device, allowing for very fast (around half an hour) analysis of DNA at the crime scene. The study examines whether crime scene investigators could work with these devices after having been given superficial training, and being supported by the research team when necessary. Interestingly, the use of the new devices did not lead to improved reasoning or better performance. This does not mean that the crime scene investigation process cannot be improved with new technology. For example, it has been suggested that better insight in the examination process can be obtained with head-mounted cameras (Baber & Butler, 2012).

To conclude, despite these studies exploring perception and cognition in a broad range of domains, one of the most neglected aspects of the crime scene investigation are perception

and cognitive processes. These, however, can be crucial for a case, as suggested above. Particularly, visual attention can select relevant perceptual input, by focusing on important places and objects. Cognitive processes using existing domain knowledge can aid the interpretation of the visual input and the guidance of visual attention, helping to select artefacts for further analysis in the lab. It is for these reasons that this thesis will focus on the perception and cognitive processes in crime scene investigation. The strategy is to compare novice and experienced CSIs and to use eye tracking technology to establish the direction of visual attention. The next chapter will provide an introduction to what is known about human perception and cognition, and will then move on describing research in other domains that examined the influence of expertise on eye movements. Based on these studies, the results in this thesis are expected to aid training and practice in crime scene investigation and to deepen the understanding of crime scene investigation process.

Chapter 2. Information Processing and Expertise

2.1. Introduction

In order to conduct a good crime scene investigation, the CSI has to assess and recognize the forensic artefacts in the scene and identify their relevance. Questions that need to be addressed are what items to sample for further analysis, how many items to collect, where to best look for these items, and what equipment best to use (De Grujter et al., 2016; Schiffer, 2009). Failing to answer these questions adequately, can lead to problems solving the case (Dror, 2015; Lee & Pagliaro, 2013). To aid the crime scene investigation process, the CSI is likely to draw upon domain specific knowledge acquired during training and upon experience from past cases (Dror, 2015). Given the types of questions that CSIs need to answer, it becomes clear that cognitive processes, such as information extraction, perception, and decision making play a key role in the crime scene investigation process.

2.2. Information Processing Approach

In cognitive psychology, a popular approach is to compare processing in the human brain to processing inside a computer (Newell & Simon, 1972). The analogue to computer processing assumes that information is processed in modules that each have their own function, and that are interconnected, as illustrated in Figure 2.1.

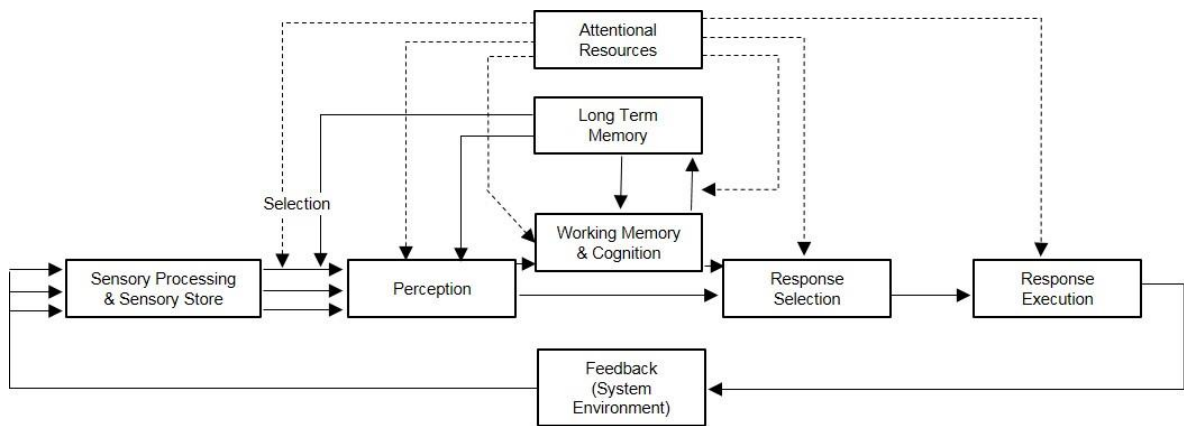


Figure 2.1. Illustration of an information processing model (adapted from Wickens & Hollands, 2000). Information is assumed to enter via senses and is processed in a range of interconnected modules for response selection and execution.

One of the first modules in this analogue deals with information selection. The information entering the senses is very rich and detailed. Processing all of this information in detail would be impossible, given the processing capacity of the brain (Feil & Mestre, 2010). For this reason, attention is selectively applied to decide which information to encode and process. The first stage of processing is known as ‘short term memory’ (STM) – also referred to as working memory (Newell & Simon, 1972). STM is assumed to be limited in capacity (Miller, 1956), and therefore, if new information is entered into STM, old information will need to be deleted first (Newell & Simon, 1972). Information in STM is then processed and transferred to long term memory (LTM) (Newell & Simon, 1972). Unlike STM, capacity of LTM has a much higher capacity. LTM is used to store general knowledge, but also experiences. Information in LTM is accessed by interaction with the STM, triggered by associated input (Ericsson & Kintsch, 1995). Strategies can be employed to make better use of the limited storage capacity in STM. One of these strategies is chunking, in which information that belongs together is grouped into larger units. An example is when being

asked to remember the letter string “I, B, M, F, B, I, C, I, A, I, R, S”, it is easier to remember it in meaningful chunks, such as “I.B.M., F.B.I., C.I.A., I.R.S.” (Gilbert, Boucher & Jemel, 2014, p.1)

Three crucial steps have been defined for an efficient and effective information processing (Wickens & Holland, 2000). The first step involves the perception of relevant information, processed with the aid of existing knowledge and experience. The second step involves processing, integrating or connecting the information in memory for interpretation and comprehension. The third and final stage involves the evaluation of the information and generating hypotheses about the situation so that an action can be selected. In these three different steps, interaction between perception and attention constrains the process, ensuring that the process does not exceed processing capacity limits.

Attention is strongly linked to eye movements, providing a way to measure it. While it is possible to shift attention without moving the eyes (as is the case in covert shifts of attention; Posner, 1980), attention shifts are often accompanied by shifts in eye gaze (Orquin & Mueller Loose, 2013). The selection of visual information by means of eye movements is known as overt visual attention (Carrasco, 2011). Because of this role for information selection, eye movement control is an essential component of human attention allocation. There is a range of eye movements that humans make. The key types of eye movements are fixations and saccades. Fixations are the periods in which the eye stays relatively still, and it is assumed that during these periods, information that enters the eye can be processed. Saccades serve to move the eye between objects and areas of interest. During saccades, the eyes rotate at a high velocity. Although information can be extracted during a saccade, typically information extraction is halted during saccades due to the excessive blur caused by the high velocity of

the eye. It is, however, still assumed that previously extracted information can be processed during a saccade (Irwin, 1998; Irwin, Carlson-Radvansky, Andrews, 1995).

The reason for making fixations, is that by fixating the eye onto a stimulus, information extraction can make use of the area of the retina specialized for detailed visual processing, known as the fovea. The remainder of the retina only provides coarse information, and while global shape and motion information can be extracted from this extra foveal area, detailed information processing from this region is limited (for example, illustrated by the phenomenon known as crowding; Levi, 2008; Whitney & Levi, 2011). The selection of where to move the eye next is known as gaze control (Henderson, Brockmole, Castelhana & Mack, 2007). This control has been extensively studied in the context of reading, where two extremes have been identified. On the one hand, the eye can be assumed to move between words in an automatic mode, and only make use of extra foveal information about where words are located. On the other hand, detailed linguistic processing can be used to establish where to look next. It is believed that actual control uses a mixture of these two strategies (Rayner, Pollatsek, Ashby & Clifton Jr, 2012).

Once researchers found methods to study eye movements, they started to gain more insight into how eye movements are used for information selection. One of the first studies into this topic by Buswell (1935) revealed that allocation of attention and selection of information is controlled by various factors (see also Tatler, 2014). These factors can be separated into low-level and high-level factors. Low level factors, leading to bottom up control of attention, are defined by visual features of the stimuli (Wolfe, 1994). A key concept is the visual saliency of objects, which is defined by a combination of factors, including colour, edge orientation, and contrast. These low level components have been incorporated into computational models

of visual saliency that predict, for a given image, what objects or areas of the image are likely to receive an observer's attention (Itti & Koch, 2000; Orquin & Mueller Loose, 2013).

High level factors, on the other hand, leading to top down control of attention, are related to the aims and goals of the observer. Together with knowledge and expectancies of the viewer, these aims and goals provide another important determinant of where attention is allocated (Henderson et al., 2007). Low level and high level factors simultaneously influence visual attention (Becker, 2014). A popular example for this joint influence is that while searching for a red object in a scene (goal related), the observer is likely to attend the other red objects in the scene that are irrelevant for the task (bottom up effects; Folk & Remington, 1998). This attention for task irrelevant features is caused by the similarity of the features of these objects with the task relevant object. Bottom-up and top-down influences, however, have a different time-course. Carmi and Itti (2006) found that visual saliency (bottom-up control) mostly affects eye movements early during visual exploration and top-down control influences attention and eye movements later on during exploration. However, in a free viewing task, Foulsham and Underwood (2007) found that visual saliency plays an active role in attention allocation during the entire trial, but not in a search task. While the interaction effect on attentional control seems strong, there is a need for further studies to address this interaction effect and explain the boundaries of top-down and bottom-up influence (Orquin & Mueller Loose, 2013).

This knowledge about eye movements can be applied to gain a better understanding of the crime scene investigation process. During collection of evidence, a CSI needs to shift attention from one object in the scene to another, and select which objects to select for further processing. This process is constrained by time factors (the scene needs to be investigated

quickly before evidence fades, but also to release the scene back to the public), emotional factors (the visual stimuli may evoke strong emotions), and by the information density of the scene (every single detail of the scene could be important to solve the crime). Information gathering during crime scene investigation is further complicated by the fact that information gathered is often uncertain, ambiguous and sometimes contradictory (Helsloot & Groenendaal, 2011). To assess the relevance of each piece of information in the scene, long term memory containing information on similar cases is accessed, and new information is interpreted in the context of this stored information. To select evidence, attention is selectively applied to an object in the scene using both top-down and bottom-up selection mechanisms (Baber & Butler, 2012; De Grujter et al., 2016; Dror, 2015). Part of the crime scene investigation process could be thought of as a visual search, but without a clearly defined search target, and requiring a continuously updated model of how different pieces of evidence fit into the case. An important component is abductive reasoning, which has the aim to find the most likely explanation of the location and state of objects in the scene (Carson, 2009). The different stages of the crime scene investigation process are illustrated in Table 2.1.

By making eye movements, the human visual system gates the stream of information entering the eyes to areas and features of the scene relevant to the task and helps to optimize the system's performance (Carrasco, 2011). As a result of the application of visual attention, relevant information can be retained (Carrasco, 2011). However, as stated above this process can also be constrained by time factors, emotional factors and information density. For example, allocation of attention and selection of information change over time in scene viewing (Unema, Pannasch, Joos, & Velichkovsky, 2005) just like the places where people look (Buswell, 1935). Likewise, studies have shown that emotional stimuli receive more

attention than neutral stimuli (Humphrey, Underwood, & Lambert, 2012), and that emotional stimuli consequently leave stronger memory traces, as in the weapon focus effect (Loftus, Loftus & Messo, 1987), where information about the weapon receives strong attention and is stored in memory, while the surrounding receive less attention and can often not be remembered. However, there has been a debate about the reason for the weapon focus, with suggestions that it might be the threat or unusualness of the object that drives the effect (for a review see Fawcett, Russell, Peace & Christie, 2013; Pickel, 1998). Most of the studies reviewed are inevitably focused on the threat dimension and scrutinizing the emotional arousal of participants. In addition, information density and distribution is at the heart of the push to develop models of salience and image information for scene viewing. Factors influencing the salience of objects include colour, edge orientation, and contrast and these low level components have been incorporated into computational models of visual saliency that predict, for a given image, what objects or areas of the image are likely to receive an observer's attention (Itti & Koch, 2000; Orquin & Mueller Loose, 2013). The selective mechanism of attention sometimes, however, poses limitations on the system, meaning that not all information in the scene can be processed. This leads to effects such as change blindness, where observers fail to notice changes in a scene over consecutive presentations (Beck, Levin & Angelone, 2007). Part of this dissertation will make use of this phenomenon, and therefore I will continue with what is known about this phenomenon.

Table 2.1.

The underlying perceptual and cognitive processes during the crime scene examination process

Task Requirement	Perceptual and/or Cognitive Process
What to look at	Visual search and identification of relevant items (artefacts)
Where to look for	Visual search and specifying the places which yield possible evidence
Interpretation and evaluation of evidential value	Interpreting and engaging the value as to why the gathering of evidence for further analysis, quality and quantity to collect

2.3. Fallibility in Perception

Rensink (2008) identified four different aspects of perception and cognition that can be studied using the change blindness phenomenon. First, change blindness can inform how the human brain detects changes. Second, change blindness will inform how attention is allocated. Third, the phenomenon can serve to understand how different elements in the scene are combined into a global perception of the scene. At the fourth and highest level, change blindness can serve to influence beliefs about how we perceive the world and act in response to it. Studies indicated that, in order to successfully perform the change blindness task, focused attention should be directed to the relevant location pre and post changes need to be encoded and compared, and the difference needs to be recognized (Simons, 2000). Short term memory plays a crucial role in this process (Beck & Levin, 2003).

Studies have revealed a range of different factors influencing change detection performance. There is an effect of scene context (Rensink, O'Regan & Clark., 1997), the need for focused attention (Henderson & Hollingworth, 1999), visual saliency of the stimuli (McCarley et al., 2004), the semantic relevance of the stimuli to the scene context (Hollingworth & Henderson, 2000), and the size of the change (Henderson & Hollingworth, 2003). In addition to these different factors, training and knowledge play a crucial role (Rensink, 2008). For example, Werner and Thies (2000) found strong expertise effects in a change blindness paradigm in American football players compared to controls, particularly for changes related to the game. The football players were faster and more accurate in locating domain-specific changes (e.g., adding another ball to the scene) compared to the controls. In contrast, no expertise effects were found in the detection of changes not meaningful for the domain (e.g., changing the colour of referee's glove or changing the location of a player's shadow) or in the detection of changes outside the domain (e.g., traffic scenes). Similar effects were found in the domain of radiology (Beck et al., 2013). As for the football players, radiology experts were more sensitive to domain-related changes, but no expertise effects were found for domain-unrelated changes. Furthermore, a recent study in problem solving in physics showed that experts detected domain related changes more quickly and accurately than controls, but no such advantage for experts was found domain unrelated changes (Morphew, Mestre, Ross, & Strand, 2015).

As explained so far, a search task like crime scene examination comprises of visual examination (perception) and comprehension and evaluation of categorical information (cognition). To deal with the limitations on the processing capacity of the cognitive system, eye movements are used to select relevant information for further processing. Because of the anatomy of the human eye, with a high density of retinal cells at a selective area and a lower

density of cells in surrounding areas, only a part of the visual scene is perceived in high spatial detail at any given time. Changes in information that is unattended, and therefore perceived with lower spatial accuracy, can therefore be difficult to detect. This is particularly demonstrated in the change-blindness task (Beck, Peterson & Angelone, 2007). Such failure in noticing changes in a scene can provide useful information in the understanding of how expertise may override limitations of the cognitive system. Change blindness results also explain why crime scene investigators may miss crucial information for a case. Expertise may overcome such change blindness effects by better selection of visual information by means of eye movements (Rensink, 2008).

Further possibly important cognitive factors involve emotional aspects. Research has suggested that expectation and motivation both strongly drive cognition and perception, particularly in forensic domains. Crime scenes tend to be highly emotional, and therefore it is important to study how such emotional contents influences perception and cognition in crime scene investigators. The approach in this thesis will be to study expertise effects on the processing of emotional (crime scene) and neutral scenes, and to compare a broad range of measures, including pupil diameter, eye movement patterns, and memory for the images.

2.4 Emotion as a factor in information processing

Traditional information processing models tend to ignore the influence of emotion on cognition. Recent studies, however, have indicated that emotion and cognitive processes are integrated and influenced by one another (see also Phelps, 2006). Cognitive processing is required to bring out emotional responses and accordingly, emotional responses modulate and lead cognition so that adaptive responses are made to the environment. (Brosch, Scherer,

Grandjean & Sandera, 2013). The interaction between cognitive process and emotions starts with emotional stimuli and then turns into emotional state which consequently leads to behavioural reactions (see Winkielman, Knutson, Paulus, & Trujillo, 2007, for a review). Behavioural, psychophysical and neuroimaging studies applying mood inductions, emotional priming, and stimuli with emotional content highlight the role of emotional stimuli and one's emotional state on cognitive processes, memory, decision-making, reasoning, and attention (Winkielman et al., 2007). The emotional state of the observer, and the emotional nature of the stimuli influences how information is stored in LTM and later utilized (Buchanan, 2007). As emotional contents are more easily manipulated in the experiments, this thesis will focus on this aspect of emotion in crime scene investigation.

Valence and arousal are the most frequently used dimensions to define emotional stimuli (Brosch et al., 2013), although in the face perception literature other dimensions, such as fear and happiness, have been considered (Angie, Connely, Waples & Kligyte, 2011). Valence provides an indication of the type of emotion experience (i.e., pleasant or unpleasant; positive or negative). Arousal provides a measure of the intensity of the emotion (e.g., low or high), independent of the valence (Winkielman et al., 2007). Both dimensions can be assumed to influence the effect on cognition, memory, and perception, attention. For example, it has been shown that emotional stimuli receive more attention than neutral stimuli (Humphrey, Underwood, & Lambert, 2012), and that emotional stimuli consequently leave stronger memory traces, as in the weapon focus effect (Loftus, Loftus & Messo, 1987), where information about the weapon receives strong attention and is stored in memory, while the surrounding receive less attention and can often not be remembered.

Evaluation of emotion differs greatly between individuals, in terms of both quality and intensity of reaction to similar experiences. For example, people can experience the same type of trauma, but not everyone will develop post-traumatic stress disorder (PTSD) symptoms (Kamphuis, Emmelkamp & Bartak, 2003; Olf, Langeland, Draijer & Gersons, 2007; Stein, Jang, Taylor, Vernon, & Livesley, 2002). The subjective evaluation of emotions depends on someone's previous experiences, goals, needs, and personality. Other factors, such as experience, knowledge and practice are likely to have an effect (Brosch et al., 2013; Lazarus & Alfert, 1964; Montagrin, Brosh & Sander, 2013).

A range of methods have been developed to measure the response to emotional stimuli. These include pupillometry (Bradley, Miccoli, Escrig, & Lang, 2008), electrodermal activity (Gavazzeni, Wiens & Fischer, 2008) or neural activity (Functional Magnetic Resonance Imaging (fMRI) Responses; Fischer et al., 2005). In addition, the qualitative aspects of emotional responses are inferred from self-reports or dedicated scales (e.g., Bradley et al., 2008). The influence of emotion on cognition has been found to depend on the goal of the task (Brosch et al., 2013).

This dissertation will use a range of methods to examine the influence of experience on emotional responses. These include pupil responses, eye movement patterns (to measure attention), subjective ratings, and memory effects. The experience groups are chosen so that they differ mostly on experience (forensic students versus students in other topics), because age, and life experience has been shown to affect the influence of emotional stimuli, although results have not always been consistent. For example, older adult's electrodermal activity was lower compared to younger participants when facing negative images. These electrodermal measurements contrasted subjective ratings, which were found to be higher in

elderly (Gavazzeni et al., 2008). Furthermore, enhanced memory in younger adults compared to older adults, was found to be restricted to negative images (Kensinger, Garoff-Eaton, & Schacter, 2007).

2.5. The Expert and Novice Paradigm

Cognitive processes have been analyzed in a range of domains (Charness & Tuffiash, 2008; Hegarty, Mayer & Monk, 1995; Hodgson, Bajwa, Owen & Kennard, 2010; Rayner, 1998; Tangen, 2013). Due to a basic interest in motor skills and the learning processes involved in these skills, research comparing experts and novices has a long history. After de Groot's (1946) research on cognitive differences between expert and novice chess players, Simon and Chase (1973) were the first to develop a theory of expertise, in which expertise is described as the use of chunking of information to encode information into abstract templates (Ericsson, 2006). The domains investigated vary from reading, visual search, problem solving, sports, driving, aviation, medicine to scrabble (e.g., Bellenkes, Wickens, & Kramer, 1997; Charness, Reingold, Pomplun, & Stampe, 2001; Crundall, 2016; Halpern & Wai, 2007; Jarodzka, Scheiter, Gerjets, & Van Gog, 2010; Krupinski et al., 2006; Kundel, Nodine, Conant & Weinstein, 2007; Savelsbergh, Williams, Van der Kamp, & Ward, 2002; Zelinsky & Sheinberg, 1995). These studies have shown that experts differ from the novices both in terms of acquired experience and in domain-specific knowledge. For example, Bellenkes et al. (1997) showed that expert pilots had superior vertical and longitudinal flight control, modulated by enhanced attention to vertical control. Experts also demonstrated shorter but more frequent fixations of the instruments, were able to adapt their viewing strategies to the task more efficiently, were better at adopting a mental model of the flight, and were found to check parameters that were constant more often. Likewise, Charness et al. (2001) found that

expert chess players could select the best move more accurately and more quickly than intermediates. Experts made larger eye movements, which were less frequent than those of intermediates. Experts fixated empty squares more often than intermediates, particularly in the first moves of the game. They also fixated relevant pieces more often than intermediates. Similarly, in judging penalty kicks from video images, expert goalkeepers were more accurate when predicting the direction of the goal kick than novices, but this could be a speed-accuracy trade-off as it was also found that experts took longer for their decisions than novices (Savelsbergh et al., 2002).

Studies, however, vary in their definition of an expert (for a discussion, see Hermens, Flin & Ahmed, 2013). Typically studies define the more experienced group as experts and the less experienced group as novices, but the expertise level of experts and novices can vary substantially across studies. Expertise is often linked to experience, but this is not necessarily an accurate measure, as not everyone acquires a skill equally fast.

Several possible definitions of “experts” have been put forward. A first definition was of experts as people with extensive knowledge and experience, influencing perception and the organization of information (McBride & Burgman, 2012). Another definition distinguished experts from novices on the basis of their skills and knowledge (Ericsson, et al., 2006). Others have suggested that experts are outliers from the population (Charness and Tuffiash, 2008, p.1), or as people with abilities allowing them outperform novices by several levels (Dror, 2011), and to deal with complex task with relative ease (Dror, 2011). Together, the definitions suggest that experts differ from novices in their levels of knowledge and skill, and allow them to perform the required tasks more efficiently. The exact levels of expertise that are used in individual studies can vary, and can depend on the availability of experts for the

study (see also the present mobile eye tracking study, where due to practical considerations, the novices were more experienced than the ones participated in desk-mounted eye tracking study).

Acquisition of knowledge does not stop after the training period and continues throughout one's professional career. As a consequence, experts can be expected to have superior perception of situations, reasoning skills, assessment and evaluation skills, leading to more accurate and efficient performance. Expertise may also lead to better planning and action (Dror, 2008). Experience can also enhance the anticipation of future actions, tasks and demands (Charness & Tuffiash, 2008). This is exemplified by the statement that experts have the ability of "seeing things differently" (Dror, 2011, p.180).

Despite these general descriptions of "experts", there are not universal standards that define expert performance, which is where the problem of assessment of skill lies (Ericsson, 1996). For example, in surgery, attempts have been made to evaluate skill on the basis of peer judgment, efficiency of instrument outcomes, or patient outcomes. The issue with many of these methods is that they are either subjective, or are linked to the difficulty of the cases presented. As more complicated cases are typically assigned to more experienced surgeons, patient outcome, for example, may not directly relate to the surgeon's skill. The only way to assess surgeons' skills would be to give cases of similar difficulty to different surgeons, but this would pose serious ethical concerns (e.g., not using the most skilled surgeons for the most difficult cases, or having less skilled surgeons perform surgeries beyond their skills). It is for these reasons that other measures such as social reputation, training and years of experience are often used to quantify an individual's level of expertise (Ericsson, 2006). Along with domain specific knowledge supported by procedural guidelines, the cases

presented to experts for training purposes or real ones shall also enrich the experts' abilities and capabilities. (Ericsson, 2006). The frequency of these cases, in other words "deliberate practices", shall lead to efficiency and effectiveness of what is being performed. The lack of a clear definition of an expert, however, makes it difficult to compare expertise effects across different studies, as one study's novices could be another study's experts (for a discussion, see Hermens et al., 2013).

Despite these problems with defining expertise, studying the effects of experience on perception and cognition can provide vital insights in what defines an expert, the different factors affecting expertise, the reasons of the errors, and provide tools for evaluation of the training programs (Tangen, 2013). A range of methods have been used to study such effects of experience on cognition and perception. An important method involves the use of verbal protocols and surveys probing into underlying mental processes. Visual attention is often studied using eye tracking, although it can also serve to gain a better understanding of decision making processes (Fiedler & Glöckner, 2012; Gidlöf, Wallin, Dewhurst, & Holmqvist, 2013; Glöckner, & Herbold, 2011). Eye tracking is not the only method used to study these cognitive processes. Others have relied on structured interviews and qualitative methods (Flin, Youngson, & Yule 2007; Patel, Arocha & Zhang, 2012). While informative, qualitative methods strongly rely on the interpretation of the researcher. As a first stage of the investigation, such methods are a valuable tool, but they should ideally be followed up by quantitative methods to study behaviour, such as with eye tracking.

The use of eye tracking technology in assessing skill has closely followed the development of the eye tracking technology in itself. Early studies of expertise often required the use of computer screens for the presentation of stimuli, while newer studies can examine expertise

effects in a more natural environment after the introduction of mobile eye tracking technology. A more detailed introduction to these techniques will be given in the next chapter.

2.6. Visual Expertise

A broad range of studies have used eye tracking to study expertise, and this thesis can therefore rely on the experience gained from these studies. The most comprehensive text in this context is probably the meta-analysis by Gegenfurtner, Lehtinen and Säljö (2011), who reviewed 296 expertise studies for performance differences across participants. Their analysis included studies of expertise in sports (badminton, squash, basketball, soccer, chess, baseball, handball, boxing, gymnastics, karate, tennis, equestrian, cricket, swimming), medicine (radiology, laparoscopy, pathology, cardiology), transportation (aviation, car driving), and other domains (biology, physics, forensics, arts, programming). Stimuli either consisted of video films, slides, schematic pictures, or eye tracking was performed in a natural environment (programming on a computer, flight simulator). Their analysis showed that most of the studies focused on how many fixations experts and novices made, what number of fixations the two groups made, on dwell times on areas of interest and saccade amplitudes. Less often measures were the time to first fixation and measures of performance not linked to eye movements, such as accuracy, and response time (Gegenfurtner et al., 2011).

In these studies, it is often assumed that knowledge acquired over time and with practice, guided by appropriate feedback and experience, attention will focus on the task relevant objects and locations for further cognitive processing (Broadbent, Horsley, Birks & Persaud, 2014; Henderson et al., 2007). What studies of attention in experts and novices demonstrate

is that novice behaviour is characterized by a stimulus driven focus of attention with a lack of a knowledge driven influence (Kurland, Gertner, Bartee, Chisholm & McQuade, 2005). This expertise-related influence is complemented by effects of general knowledge about the environment (Henderson, 2007; Henderson et al., 2007). For example, when searching for a pillow in a scene, viewers most likely try and find the pillow on the bed. Overall, Gegenfurtner et al. (2011) found out that eye movements differ across experts and novice and proposed that eye movements can be used to measure the visual expertise in various domains (see also Hermens, et al., 2013 for the domain of surgery). While some of the studies have applied a natural environment (e.g., Khan et al., 2012; Bellenkes et al., 1997), many studies have used rather impoverished versions of the actual task or skill, by relying on video clips and image.

With the development of mobile eye trackers, it can be expected that in future years, a shift will be made towards the actual skill environment, despite the difficulties in analyzing the data (this will be described in more detail in the methodology chapter). I now address individual studies to explain briefly what aspects of eye movements are different between experts and novices.

2.6.1. Sports

Expertise effects in sports have often been examined with the aim to develop methods to improve skill. As discussed, a broad range of sports have been examined for expertise effects (Gegenfurtner et al., 2001; Mann, Williams, Ward & Janelle, 2007). These studies have often made use of video based simulations, and only recently they have moved towards simulated tasks, possibly influenced by recent developments in eye tracking technology. Eye

movements in sport related tasks were often found to be context and task dependent, and were often influenced by the study design (Dicks, Button, & Davids, 2010; Mann et al., 2007). In terms of eye movements, skilled expert performers were often found to make fewer fixations with longer durations, to make better use of peripheral vision (however see also Williams & Davids, 1998; Williams, Davids, Burwitz & Williams, 1994). Some of the studies found that while executing an action (e.g. hitting the ball in golf, throwing the ball to the basket in basketball) experts displayed a longer “Quiet Eye” fixation, being the last fixation before an aiming action (Vickers, 1996; Vickers, 2016). Experts also found longer quiet eye durations than novices (Mann et al., 2007). Some studies have used this finding for the purpose of training and have demonstrated faster skill acquisition (Harle & Vickers, 2001; Wood & Wilson, 2011; but see Afonso, Garganta, McRobert, Williams & Mesquita, 2012). This expertise has generally been found to be domain specific, and attention was not found to be different between experts and novices in basic attention tasks unrelated to the domain (Memmert, Simons & Grimme, 2009; Overney, Blanke & Herzog, 2008). Performance-wise, they were better at anticipating the next moves of the opponents or the trajectory of the ball. They showed faster responses than the novice counterparts, and were better at adapting their strategies to the task (Piras, Lobietti & Squatrito, 2014; Piras, Pierantozzi & Squatrito, 2014; Savelsbergh, et al., 2002).

2.6.2. Chess

When presented with actual chess setups, expert chess players demonstrated superior memory skills compared to intermediate or novice chess players (Duchowski, 2007). Other studies showed differences in the visual span and visual perception of world class expert chess players. Using a change blindness paradigm, Reingold, Charness, Pomplun and Stampe

(2001) demonstrated a greater visual span in experts. Eye movements demonstrated that experts attended salient and meaningful locations more often. Systematic eye movement patterns were only found for meaningful configurations in a systematic way.

2.6.3. Aviation

Aviation is a well-studied domain, because of the high stakes involved. Plane crashes often cause many victims, and create bad publicity for the airline involved. Studies in aviation have examined performance differences, attention allocation, situational awareness, mental workload and stress management in pilots in various different tasks (e.g., landing, low altitude flight, target pursuit) in different aircraft, including helicopters and planes (e.g., Bellenkes, et al., 1997; Kirby, Kennedy & Yand, 2014; Schriver, Morrow, Wickens & Tailleur, 2008). Studies often employ sophisticated flight simulators and eye tracking techniques to measure attention and inattention. The findings support the view that experts are more successful, because they are more flexible than the novices in completing the task (Bellenkes et al., 1997). For example, by using a flight simulator, Schriver et al. (2008) found that expert pilots attended more informative parts of the simulator than the novices. These results were subsequently adopted to train novices to use similar scanning strategies. Likewise, the design of the cockpit was adjusted to incorporate the findings (Charness & Tuffiash, 2008). Similarly, Kirby et al. (2014) found that enhanced attention to the instrument display led to improved flight control. Expert pilots also demonstrated smaller pupil size distributions, suggestive of a lower experienced workload (Yu, Wang, Li, Braithwaite & Greaves, 2016). Similar studies were conducted in air traffic controllers. For example, Van Meeuwen et al. (2014) found that novices focused more often on the irrelevant areas, and applied inefficient strategies compared to experts. Experts were found to apply more selective

perceptual strategies, looking more often at crucial objects and targets in fewer fixations (Hyun et al., 2006).

2.6.4. Problem Solving

Another domain in which eye tracking has been used to study expertise effects is problem solving (see Spivey & Dale, 2011 for a review). For example, training in physics is often considered to be difficult. Studies of how problems are solved in physics and in similar domains revealed that, for example, when solving electrical circuit problems, experts were found to repeatedly look back at the circuit, while novices did not show such a strategy (Rosengrant, Thomson & Mzoughi, 2009). Similarly, Carmichael, Larson, Gire, Loschkyand and Rebello (2010) found that experts allocated their visual attention to thematically relevant areas of physics diagram, while novices were more focused on the salient features. Furthermore, Hegarty et al. (1995) demonstrated that students with superior problem solving performance in mathematics word problems distributed their attention more efficiently. In a related study, correct solvers of Tower of London problems showed more efficient eye movement patterns on critical locations of the problem (Hodgson et al., 2000). Three different stages in visual strategies were found. In the first (comprehension) and last stage (confirmation), the fixations were allocated on the target location, while in the second stage the possible solutions were fixated (evaluation). Increased complexity of the problems led to more time spent in the second stage. For incorrect answers, eye movement patterns were similar to those for the previous problems suggesting difficulties in flexibly shifting between strategies (Hodgson et al., 2000).

Problem solving in the context of internet use has been studied to determine how expertise influences engagement (Hyöna, 2010). In the context of multimedia learning, Jarodzka et al. (2010) found that experts more strongly focused on relevant features of the stimuli to solve the problem. They were also found to be more efficient at using non-analytical and knowledge-based shortcuts. The increased attendance in experts on relevant areas persisted, even when novices were provided with the domain-specific knowledge. Similar results in the context of paediatric neurology were obtained by Balslev et al. (2012).

2.6.5. Medicine

Another field with high stakes is medicine, where life and death decisions often have to be made under high stress conditions. The majority of the studies in medicine have focused on radiology (Beard, Johnston, Toki & Wilcox, 1990; Cooper, Gale, Darker, Toms & Saada, 2009) and surgery (Hermens et al., 2013). Studies on radiology often examine search strategies involved in detecting tumours in radiographic images. In contrast, the work on surgery focuses on the eye-hand coordination involved in key-hole surgery, a skill often difficult to acquire (e.g., Krupinski et al., 2006; Kundel et al., 2007). Other studies have examined a concept called “situational awareness”, which reflects how aware the medic is of the situation and of events that possibly indicate issues (e.g., Koh, Park, Wickens, Ong & Chia, 2011). In radiology, it was found that experts attend to the task relevant areas (e.g., abnormalities in an X-ray) faster, made fewer fixations and have larger visual spans than novices and have more efficient search strategies (Reingold & Sheridan, 2011). Studies also showed that eye tracking technologies can be used to improve the performances of novices, applying the technology as a training tool (Donovan, Manning & Crawford, 2008; see Tien et al., 2014, for a review). In surgery, evidence for more efficient looking ahead was found

in experts and training novices to adopt this viewing strategy led to better surgical performance in a simulator task (Wilson et al, 2010; see Hermens et al., 2013 for a review). Studies have also shown that eye movements can be used to classify surgeons in novices or experts (Ahmidi et al. 2010; Ahmidi, Ishii, Fichtinger, Gallia & Hager, 2012; Nicolaou, James, Darzi & Yang, 2004; Richstone et al., 2010), suggesting that the measurements can be used as a method to assess skill more objectively than can be done with peer assessment.

2.6.6. Driving

Accidents in driving causes large numbers of victims every day. Although they are not reported on in the same way as aviation accidents, accident prevention in driving is an important area of interest. Studies of attention in driving have examined the role of road types (Crundall, Underwood & Chapman, 1999) and attention and situational awareness during hazardous situations (Crundall, 2016; Chapman & Underwood, 1998; Crundall et al., 2012) and when engaged with a secondary task (using cell phones while driving; Chisholm, Caird, Teteris, Lockhart & Smiley, 2006). Driving studies have not only focused on cars, but also other vehicles such as motorcycles (e.g. Di Stasi, Contreras, Cándido, Cañas & Catena, 2011; Hosking, Liu & Bayly, 2010). The studies on driving generally make use of simulator based or video based tasks. Eye tracking in such tasks suggest more efficient visual scanning strategies in experts, that cover larger horizontally areas, although the results depend on driving conditions. Studies have examined how expert driver's scanning techniques may be transferred to novices to prevent accidents (Chapman & Underwood, 1998; see also Duchowski, 2007, for an overview).

Before turning to forensic sciences, the above literature suggests similarities of expertise effects on eye movement patterns across research domains, despite some contradictory findings. The contradictory findings show that eye movement behaviours can be domain specific and task dependent. Developments in eye tracking technology are expected to enhance the insight to more domains and to reveal expertise effects in more real-life situations.

2.6.7. Forensic Sciences

The present dissertation will apply eye tracking technology to understand expertise effects in forensic sciences. This dissertation is not the first to study eye movements in the context of forensic sciences. Two main streams of research can be identified: “forensic process in investigations” and “court appearance of the eye witnesses”. Studies in the first category examine the role of eye movements in collecting physical evidence in a crime scene. Studies in the second group examine the factors affecting the eye-witnesses’ testimonials. Both domains will be discussed in more detail below.

2.6.7.1. Studies on Forensic Lab Examinations

To better understand what skills are acquired when moving from a novice skill level to an expert skill level in crime scene investigation, studies have compared the visual attention of novice and expert CSIs during inspection of a scene. This method closely matches that used in studying expertise effects in surgery (see Hermens et al., 2013 for a review). The first studies in the domain of crime scene investigation involved inspections of handwriting and fingerprint examinations. The processes and strategies used in these domains show

similarities with radiology, and studies could therefore rely on the methods and technique to study expertise effects in this domain. Comparison based examinations in the forensic laboratories have been criticised, with critics putting forward a lack of scientific validity and reliability procedures (NAS Report, 2009). In the studies, examiners were asked to compare samples of fingerprints or signatures, either on a computer or on paper. An overview of the studies in this domain is provided in Table 2.2.

All studies compared experts and novices. Studies either used real (Dror & Charlton, 2006) or simulated cases, where in the latter case, participants often knew that the case was simulated. The very first study to use eye tracking in forensics, examined how forensic examiners (experts) and students (novices) examine signatures (Dyer, Found & Rogers, 2006), using a Tobii 1750 eye-tracking system (Tobii Technology, Sweden). The task of the participants was to determine whether the signatures were genuine, disguised or forged. Eye movements (fixations, scan paths and dwell times) suggested that both groups used almost the same strategy, which was to look at the salient features of the signature. Despite these similarities in viewing patterns, the experts outperformed the lay people, suggesting that eye movements are driven by bottom-up features, while performance is in part drive by top-down effects. This was later interpreted as an enhanced capacity to process local information aided by more global information (Dyer, Found & Rogers, 2008).

A second study also compared forensic document examiners (experts) and students (novices) with the same equipment (Tobii 1750 system). This second study found reaction time differences between the groups, with longer examination times for experts. Experts were also more cautious in their decisions than novices, suggesting that experts prefer to say that they

are no certain rather than making errors, consistent with research by Sita, Found and Rogers (2002).

Table 2.2.

Studies that used eye tracking technology in traditional forensic disciplines comparing the domain of research and the eye tracking measures used.

Domain	Eye Tracking Measures	Authors
Forensic Signature Analysis	Fixation Durations on AOIs	Dyer et al., 2006
Forensic Signature Analysis	Fixation Durations on AOIs	Dyer et al., 2008
Forensic Signature Analysis	Fixation Durations on AOIs	Merlino, 2014
Forensic Fingerprint Analysis	Saccades, Fixation Durations, Number of Fixations	Busey et al., 2011
Forensic Fingerprint Analysis	Saccades, Fixation Durations, Number of Fixations	Busey, Yu, Wyatte & Vanderkolk, 2013
Forensic Fingerprint Analysis	Saccades, Fixation Durations	Allinson & Adams, 2012
Crime Scene Examination-Evidence Search	Process Tracing - Observations from records of data from cameras	Schraagen & Leijenhorst, 2001
Crime Scene Examination-Evidence Search	Process Tracing - Recordings of data from the head mounted point of view camera	Baber & Butler, 2012

More recent work on forensic document examinations (Merlino et al., 2014) examined how strategies vary across experts and novices, now using a Tobii T-60 system. A much larger sample of experts and novices took part than before. Results replicated earlier finds showing that experts outperformed novices. The larger sample size also revealed differences in how the tasks were approached, with a more focused approach of experts.

Similar studies were conducted in the domain of fingerprint examinations (e.g., Busey et al., 2011; Busey et al., 2013), although eye tracking in this domain has not been as popular as in other domains. In the first study of its kind on eye movements in fingerprint examinations, Busey et al. (2011), using a 1750 Tobii eye tracking system and a mobile eye tracking system (Positive Science, USA), did not detect any differences in average fixation durations between groups. They, however, did find that experts spend more time to look at the latent prints and made shorter saccades, possibly by a stronger focus on the informative parts of the latent prints and by chunking features for memory consolidation. When time was unlimited, experts showed large variability in how they approached the fingerprint examinations.

In a re-analysis of their original data, Busey et al. (2013) tested the usability of an algorithm to classify eye movements in fingerprint examinations. Again, no differences in fixation durations or numbers of fixations were found, and again, experts made shorter saccades than novices. The algorithm suggested that experts found more matching regions than the novices.

In an unpublished study by Allison and Adams (2012), using a Tobii 1750 eye tracker, no differences in eye tracking measures such as the number of fixations and fixation durations were found between novices and experts, both in a computer based task, and a more realistic setting. Interestingly, a change blindness paradigm revealed longer fixation durations for

experts, without any differences in detection rates across the two groups. This study confirmed earlier observations suggesting that experts are more consistently looking at the crucial areas. In this case, novice participants spent more time than examiners.

2.6.7.2. Eye-Witness Work as a Part of Criminal Investigations

Eyewitnesses play an important role in cases, providing important insights about perpetrators, and the order of events in a crime. Because of the importance of this eyewitness evidence, the factors that affect the eye-witnesses accuracy and decision making process have been studied extensively (e.g., Flowe & Cottrell, 2010; Hope & Wright, 2008; Josephson & Holmes, 2011; Loftus et al., 1987; Mansour, Lindsay, Brewer & Munhall 2009; Megreya, Bindeman, Havard & Burton, 2012). The first study to examine the role of visual attention, by Loftus et al. (1987), examined the weapon focus effect. Participants' eye movements were recorded while they watched a video about a customer in a restaurant holding either a gun or a check. Longer fixation durations on the gun than on the check were found, providing a direct measure of the gun weapon focus effect. Mansour et al. (2009) used eye tracking to examine decision making strategies in simultaneous line-up identifications. Interestingly, participants' eye movements were not a good predictor of the accuracy of the decisions. Participants spent more time comparing the faces in a line-up when they did not make a decision than when a suspect was selected. In agreement with these findings, Flowe and Cottrell (2010) found that longer response times in trials without a decision. When a suspect was selected, the face of this suspect was fixated for longer, and particularly the first fixation on the face of the suspect was longer. Incorrect positive identifications, in contrast were associated with more repeat visits on the face and longer viewing durations.

Josephson and Holmes (2011) conducted a study on the effect of race in eyewitness line-up identification. In their study, participants (Anglo and African Americans) were asked to watch a crime video. After a 24-hour interval they were asked to identify suspects from a line-up. Two videos were used: One in which the perpetrator was an Anglo-American and another with an African American perpetrator. The results showed that participants were less accurate when the perpetrator was of the opposite race than when the perpetrator was of the same race. They also found that Anglo-Americans tended to make quick decisions with fewer eye fixations and photo-by-photo comparisons than African-Americans who were more cautious.

Megreya et al. (2012) examined whether the bias towards suspects towards the left of the line-up can be explained from eye movement patterns. Their eye movement results showed a strong left-to-right scanning bias, explaining the decision bias.

2.7. Conclusion

Studies differ in how to distinguish between experts and novices. They also vary in the number of levels of expertise considered (novices versus experts, or novice, or with an intermediate level of expertise; Gegenfurtner et al., 2011) or whether expertise is considered on a continuous scale (e.g., measured as the number of years of experience). Expertise can be defined by performance, years of experience, social reputation, and level of training (Hermens et al., 2013). The lack of consistent definition of experts makes the comparison across studies difficult. The consequences particularly arise when eye movements are to be used for assessment, because if expertise for the training set data cannot be uniquely defined, this will also lead to problems predicting expertise on the basis of eye movements (Hermens

et al., 2013). Likewise, using eye movements as a measure of the development of skill depends on a clear definition of expertise in the data used to create the scale of measurement.

Studies also vary extensively in the context in which eye movement are measured. Some of such difference may be crucial (e.g., between watching and performing surgery, Hermens et al., 2013), whereas others (e.g., between actual and simulated surgery) lead to smaller eye movement differences (Atkins, Jiang, Tien & Zheng, 2012; Tien, Atkins & Zheng, 2012). It is important to establish which variations make a difference and which are less important and therefore the studies presented here will examine expertise effects across a broad range of tasks (e.g., memory, change detection, exploration) and situations (watching crime scenes, versus actively exploring). Lab studies have the advantage of stronger control about the stimulus conditions, whereas real-life studies have higher ecological validity (Dror & Cole, 2010).

By changing the conditions of study across experiments, the advantages of different approaches are combined. Two of the tasks examine eye movements during crime scene investigation. This was done either on a computer, where participants viewed images of crime scenes presented on a computer screen and were asked to click on possible evidence, or while participants actively engaged with a simulated crime scene, wearing a mobile eye tracker to record their eye movements. The first setup has the advantage that eye movements can be precisely recorded, at high spatial and temporal resolution, and that the visual input is identical across participants. The disadvantage of the approach is that participants only see images taken from one point of view, and that their head movements are restricted to improve accuracy of the eye tracking. The second setup has the advantage of being much more like actual crime scene investigation, but has the disadvantage that eye movement recordings are

less precise, that the visual input is no longer equal across participants (participants can move their heads and bodies around, so not every part of the scene is in view for the same amount of time for each participant), and that analysis of the eye movements is much more time-consuming and more subjective (it depends on how the coder assigns eye movements to specific regions of interest). The on-screen task is complemented by two further on-screen tasks. The first one makes use of past work suggesting that expertise can modulate how easy it is to find changes in a scene, but only for scenes relevant to the task in which the person has expertise. The second task examines the emotional aspects of crime scene investigation, and aims to establish whether experience with crime scene investigations influence the emotional response to crime scenes.

Chapter 3. Methods

3.1. Introduction

Cognitive psychologists have used a range of methods to study mental processes. These include behavioural measures, such as reaction times to stimuli (Townsend & Ashby, 1983; Luce, 1986), accuracy of responses, such as in psychophysics (Swets, 2014), hand or mouse trajectories (Freemann & Ambady, 2010; Freeman, Dale, & Farmer, 2011; Hehman, Stoller, & Freeman, 2015), neuroimaging (Parris, Kuhn, Mizon, Benattayallah, & Hodgson, 2009), neuropsychology (Snyder, Nussbaum, & Robins, 2006), qualitative methods (Denzin, & Lincoln, 1994) and electrophysiological measures, such as heart rate, skin response, pupil dilation (e.g. Ahlstrom, & Friedman-Berg, 2006). I here focus on another class of methods which make use of the movements of the eyes, as revealed by eye tracking. In the following, I will start with a brief history of eye tracking, followed by a description of the types of eye movements that people make, what measures of eye movements are commonly used, and end with a discussion of the technical aspects of eye tracking.

3.2. History of eye tracking

Keith Rayner, in his seminal reviews of the eye tracking literature (Rayner 1978; Rayner, 1998), described three eras of eye movement research. The first era started with the first application of eye tracking in the psychology of reading in 1879, and ended around 1920. In this era the most basic elements of eye movements were discovered, such as the distinction between saccades (fast eye movements) and fixations (periods in which the eye remains -

relatively- still) and the idea of saccadic suppression (that no information is normally extracted during saccades). In this era, also the use of saccadic latencies as a measure of cognitive processing and the concept of visual span were established.

The first era was followed by the second era, which spanned from around 1920 to around 1950, and which had a rather applied focus. Eye movements in scene perception was an important field of study in this era, for example, in the work by Buswell (1935), who found evidence for bottom up and top-down factors in the control of eye movements. At the end of the second era, there was a gap until the third era, as defined by Rayner (1978; Rayner,1998), as there was a belief that everything that could be learned about eye movements had already been discovered.

The third era, starting in the mid-1970s, was sparked by developments in eye tracking technology (Jacop & Karn, 2003; Wade & Tatler, 2005). These developments were driven by developments in computer technology, allowing for the processing of large amounts of data, so that more sense could be made of eye tracking data than initially was possible. For example, in the work by Buswell (1935) and Yarbus (1967), analysis mainly focused on visual inspection of the scanpaths. Using computer technology, more advanced data analysis became possible, analyzing eye movements for gaze durations, saccade amplitudes, and refixation frequencies, such as common in research on eye movements during reading (Rayner et al., 2012). The use of computers also allowed for changing the stimuli while people were scanning an image. This was extensively used in research on how the eyes move during reading of text. For example, by changing letters beyond a window around fixation in a text by either jumbled letters, XXXs with or without spaces, it could be examined what

information is taken in from beyond fixation, and how large the visual span around the fixated word is (McConkie, 1997; McConkie & Rayner, 1975; Rayner, 1975; Reder, 1973).

It is also in the third era that theories started to be developed about the where and when of eye movements (Engbert, Nuthmann, Richter, & Kliegl, 2005; Findlay & Walker, 1999; Reichle, Rayner, & Pollatsek, 2003; Trappenberg, Dorris, Munoz, & Klein, 2001). Before the introduction of these more complex theories, more basic hypotheses were formed, such as the “eye-mind hypothesis” by Just and Carpenter (1980), which stated that what is being fixated on is what is being processed (with a minimal lag). Following this hypothesis, fixation durations became the key measure of processing of a visual stimulus: The longer an object, or area of the scene was fixated on, the more thoroughly it was processed.

The eye-mind hypothesis has been challenged, for example by the observation of covert attention (Posner, 1980), and more specific tests followed (Anderson, Bothell, & Douglass, 2004; Fox, Merwin, Marsh, McConkie, & Kramer, 1996). Despite these possible violations of the assumption, a common assumption is still that what observers are looking at is what they are currently processing.

Key areas of eye movement research have been eye movements during reading (Just & Carpenter, 1987; Gibson & Levin, 1975; Rayner et al. 2012) and eye movements in scene perception (Henderson, 2003; Irwin & Zelinsky, 2002). Particularly, in the domain of reading, a thorough understanding of the control of eye movements was obtained. This understanding was implemented in detailed models of eye movements (Engbert et al. 2005; Reichle et al., 2003) and made use of sophisticated techniques to study what exactly is

processed in extra foveal vision, and what determines fixation durations and saccade amplitudes (Rayner et al., 2012).

More recent developments, triggered by mobile eye trackers becoming commercially available, focus on eye movements in day to day tasks, including tea making (Land, Mennie, & Rusted, 1999), walking (Foulsham, Walker, & Kingstone, 2011) and social interaction (Ho, Foulsham, & Kingstone, 2015; Macdonald, & Tatler, 2013). Other more recent eye tracking applications involve understanding how observers watch video or other forms of multimedia (Jarodzka et al. 2010), how problem solving works (Carmichael et al., 2010; Hodgson et al., 2000), and how people make decisions (Hegarty, Canham & Fabrikant, 2010; Stewart, Hermens & Matthews, 2016).

3.3. Types of Eye Movements

People make a range of eye movements, each serving a different purpose. In this section, I will briefly describe the most important of them and how they relate to behaviour.

3.3.1. Fixations and saccades

Two key types of eye movements, particularly for head-mounted eye tracking, are fixations and saccades. Although the fixations are not eye movements and may be considered as the “phases” of eye movements, it is assumed that the majority of visual information extraction occurs during fixations. Fixations are the moments during which the eye stays relatively still and alternated by saccades, which are fast, often ballistic eye movements that bring the eye from one fixation to the next. It is assumed that no information extraction takes place during

saccades. Fixations and saccades are the key eye movements observed during tasks such as reading, visual search, and free viewing of images, particularly if the head is restrained during eye tracking.

3.3.2. Fixational eye movements

These are eye movements which take place during periods of fixations. They are relatively small amplitude movements, and are assumed not to interfere with information extraction, although a case has been made for the involvement of fixational eye movements in covert attention (Engbert, & Kliegl, 2003; Laubrock, Engbert, & Kliegl, 2005). Three types of fixational eye movements have been distinguished, each with their own properties. These are tremor, small amplitude, seemingly random movements of the eye, and seemingly controlled for each eye independently. Second, there is the slow drift movement, in which the gaze position of the eyes slowly drifts away from current fixation. Finally, there are small eye movements that resemble saccades in almost every aspect, except that they occur when observers try not to move their eyes and have much smaller amplitudes which are known as microsaccades (for reviews, see Martinez-Conde, Macknik, Troncoso, & Hubel, 2009.; Rolfs, 2009). While tremor and oculomotor drift are difficult to measure with most types of eye trackers, microsaccades can be detected with a range of eye trackers, and special detection algorithms have been introduced to pick them up from signals recorded with the more common video based eye trackers (Engbert, & Kliegl, 2003; Engbert, & Mergenthaler, 2006).

3.3.3. Smooth pursuit and Opto-Kinetic Nystagmus (OKN)

When dealing with moving stimuli (even while keeping one's head still), two further types of eye movements can be observed. While smooth pursuit occurs when tracking a single target, OKN is found when the entire surrounding of the observer moves (e.g., when the observer is placed inside a rotating drum). Both types of eye movements serve to stabilize the external world on the retina, whereas in smooth pursuit the aim is to stabilize the foveated object, whereas in OKN the aim is to stabilize the entire visual field. Another difference between the two types of eye movements is that smooth pursuit eye movements are typically observed for relatively small movements of a visual target. When the target movements are larger or when the tracked object occupies a larger part of the visual field, pursuit changes into OKN, also known as railway nystagmus (Carpenter, 1988; Wade & Tatler, 2005). OKN typically involves saccades, which return the eyes to their initial position. In a clinical setting, recent studies have suggested that smooth pursuit deficits may be an important tool in the diagnosis of schizophrenia (Benson et al., 2012), and it has been suggested that such application may be extended to other conditions, such as major depressive disorder.

3.3.4. Vestibular oculomotor reflex (VOR)

VOR occurs when the head is allowed to move, and will therefore play a limited role during fixed head eye tracking, but will come into play during mobile eye tracking. VOR compensates for the rotation of the head, ensuring that stimuli remain stable in the retinal image, preventing blur that could interfere with processing of the retinal image (Wade & Tatler, 2005). VOR makes use from the signals of the inner ear (the semi-circular canals),

and from visual input. If the two are misaligned, this may lead to motion sickness and vertigo, and making the testing of the VOR an important clinical assessment tool.

3.3.5. Vergence eye movements

The final type of eye movements is vergence eye movements. Vergence eye movements ensure that the same object is fixated by both eyes. These eye movements typically occur when fixating a looming or receding object, or during self-locomotion, while moving forward or backward. In eye tracking studies, vergence is not often incorporated, either because only one eye is being tracked, or because two-dimensional stimuli are used (e.g., images presented on a computer screen). In 3D scenes, and mobile eye tracking, vergence becomes an important clue as to which object is being fixated, although similar information may often be extracted from object occlusions.

3.4. Most Common Eye Movement Measures

Early studies of eye movements (Buswell, 1935; Yarbus, 1967) focused mostly on the visual inspection of scanpaths. These studies often made use of few participants (partly because of the obtrusive nature of the eye tracking methods used), which made such methods feasible. However, when trying to draw conclusions on the basis of data of more participants, aiming to obtain a measure of eye movements across the population, it is important to quantify the observed eye movement patterns in some way or form. This extraction of measures was facilitated by the introduction of computers (Rayner, 1998), and computer programming languages, allowing individual researchers to develop their own algorithms for analyzing eye movement data, such as found recently in methods to compare scanpaths (e.g., Cristino,

Mathôt, Theeuwes, & Gilchrist, 2010), with or without a delay, as in cross-recurrence analysis (e.g., Richardson, & Dale, 2005), and to develop models to summarize observed eye movement patterns (e.g., Nuthmann, Smith, Engbert, & Henderson, 2010).

An eye movement analysis often starts with the analysis of basic parameters of the two fundamental types of eye movements: Fixations and saccades. What we know about the link between these eye movement measures and cognition is often based on research in reading (Rayner et al., 2012), where very controlled experiments have been carried out to reveal the link between understanding of the text, complexity of the input and the various eye movement measures. Due to the richness of measures provided by eye movements, there is no fixed method for analyzing and interpreting eye movement data (Jacop & Karn, 2003). One proposed method is to combine think aloud protocols with eye tracking. Eye tracking data in such protocols are used either to aid the think aloud process (retrospective think aloud) or to provide additional information while participants think aloud (concurrent think aloud). (Cooke & Cuddihy, 2005; Elling, Lentz & de Jong, 2011; Gerjets, Kammerer & Werner, 2011; Hyrskykari, Ovaska, Majaranta, Rähkä & Lehtinen, 2008)

The approach taken here is to combine more traditional eye movement methods, such as dwell times, fixation durations, and saccade amplitudes, which may provide information about how long participants spend fixating certain objects or parts of the scene, the overall speed by which the image or scene is scanned and the extent to which the eye jumps while scanning the scene, with verbalizations from the participants about their thought processes. I will first discuss the more traditional methods, before explaining the retrospective think aloud protocol in more detail.

3.4.1. Dwell time

This is the total time spent on an area of the stimulus (Jacop & Karn, 2003). In the study of eye movements of reading, this time is often referred to the total fixation time, and it is assumed to reflect the amount of processing needed to process the stimulus. In reading, it has been found that words are often refixated by means of regressive eye movements (eye movements taking the eye back to an already processed word in the text) when readers appear to reinterpret the sentence, as in garden-path sentences (Rayner et al., 2012). By taking into account both periods of fixation (initial fixation and refixation), the total processing time of the word is better measured than when considering the first fixation only. In reading, it has also been found that longer words are often fixated multiple times. Part of the reason, may be that the initial fixation on the word (often aimed at the center of the word) is not sufficient for processing longer words accurately (due to the visual span). Another reason is that longer words are often less frequent and more complex and require longer processing. By pooling these various fixations of the same word into one measure, this overall processing time is better taken into account (Rayner et al., 2012). Dwell times may be measured as the overall time (i.e., in milliseconds), or the relative time (i.e., as a percentage of the overall viewing time of the image or the scene).

3.4.2. Number of fixations

As indicated, it is not uncommon that observers fixate the same object or the same stimulus multiple times, which can be to improve detailed processing (if the first fixation was at a suboptimal position), or to process the stimulus further after moving away from the stimulus initially. To measure this further processing, it is therefore useful to measure the number of

fixations on an object or area of interest (Jacop & Karn, 2003). While the number of fixations is often measured as a single number (for both within and between object refixations), it would be possible to split the number of fixations into those from saccades within the same object, and those from saccades between objects, as they may reflect different reasons for having another (or longer) look at the object. Often, when plotted against each other in a scatter-plot, quite a strong association is found between the overall dwell time on an object and the number of fixations towards the object. This is because the eye has the tendency to make a saccade after a relatively fixed interval (although variations are possible due to, for example, differences in text complexity; Rayner et al., 2012), which has sometimes been taken as an indicator for an automatic generator for eye movements, producing new eye movements at a relatively fixed interval.

The number of fixations are most commonly measured for a particular region of the screen, defined as “areas of interest” (AOI), rather than for the entire screen. These areas can be objects, parts of the scene (e.g., the background), words, search objects, and basically reflect what the researcher thinks are important areas for the task at hand. An alternative method is to define AOIs using the observed eye movements or verbal or manual responses by the participants, known as data driven AOIs. For example, Santella and DeCarlo (2004) developed an automatic algorithm based on a mean-shift clustering technique using the participant’s fixation data. The drawback of this method, however, is that it defines different AOIs for each participant, and therefore cannot be used to generalize across participants. A similar approach was taken by Stewart et al. (2016), who clustered the eye movements towards pairs of gambles (e.g., a 50% chance of £50) using a k-means clustering algorithm (assuming five clusters: one for each attribute, and one around fixation). When fixations to each of the clusters were compared to fixations classified by traditional AOIs, similar, but

not identical results were obtained. In this dissertation, I will use a similar approach, but then use areas around objects mentioned in the verbal reports of the participants or around mouse-clicked objects.

3.4.3. Average fixation duration

A more global measure is the average fixation duration, which may relate to the complexity of the task. For example, in reading it has been suggested that average fixation durations are longer when the text is more complex (Rayner et al., 2012), as is for example the case for mathematics books compared to tabloid newspapers. Longer average fixations could also indicate more thorough processing overall, or a more parsimonious scanning technique, avoiding a large number of saccades.

3.4.4. Time to first fixation

The time to first fixation is a measure that is more informative for displays that are not typically approached with a standard scanning pattern (i.e., time to first fixation may be less relevant for fixations on words in texts, as these are normally fixated in the order in which they are presented). It refers to the time between the onset of the stimulus, until the first fixation, and may reflect how salient or relevant the fixated object is for the task. Time to first fixation may also reflect the efficiency of finding objects, such as in visual search, and experts may be more efficient at locating relevant objects and may therefore have shorter time to first fixations.

3.4.5. Saccade amplitude

The size of the jump that the eye makes between fixations, the saccade amplitude, may provide information about the scanning strategy of participants. In reading, for example, larger saccades may be expected (more word skipping) for more skilled readers, and for texts that are easier to understand. In inspecting a scene, a more global (or holistic) scanning strategy may be reflected by larger saccade amplitudes, while more detailed scanning may result in smaller saccade amplitudes. Histograms of saccade amplitudes across a sampling interval tend to be skewed, with a large frequency of small amplitude saccades, and it is therefore important not to just consider average saccade amplitudes, but also other measures such as the median saccade amplitude, to better reflect this distribution.

3.4.6. Saccadic latencies

Saccadic latencies are the time between the start of an eye movement task and the execution of the correct saccade. They are found, for example, in the anti-saccade task, in which participants are presented with a peripherally presented target and are asked to make an eye movement in the direction opposite to this target, as quickly and as accurately as possible (Munoz & Everling, 2004). Saccade latencies thereby indicate how fast a stimulus can be processed or a task be completed, and are less contaminated by noise in the time required for preparation of the motor command to execute the response, compared to manual responses. Saccadic latencies are therefore often faster, and less variable than manual responses (e.g., Friesen & Kingstone, 2003).

3.4.7. Scanpath measures

The measures above only deal with individual fixations and saccades. There are techniques to compare the sequences of fixations (e.g., Cristino et al., 2010). The overlap between scanpaths can be expressed as a measure, and then be compared across groups. The analysis of scanpaths is still in development, but in this thesis I will apply the most recent of these developments, cross recurrence analysis (Richardson & Dale, 2005), which can also take into account similar patterns, but occurring at a delay.

3.4.8. Blinks

Blinks serve to spread the tears across the eye, preventing it from drying out. It is estimated that people blink on average 17 times per minute, but the rate of blinking has been found to vary with task, with a 55% reduction while reading (compared to rest) and an almost 100% increase during social interactions (Bentivoglio et al., 1997). During blinks, no information can be extracted, and it is assumed that blinks are also related to a disengagement of attention (e.g. Nakano, Kato, Morito, Itoi, & Kitazawa, 2013). Studies have suggested that blinking is a sign of fatigue and mental workload (e.g., in aviation Ahlstrom, & Friedman-Berg, 2006; in driving Benedetto et al., 2011; in surgery Zheng et al., 2012), and that blink durations and blink rates decrease when the workload increases.

3.4.9. Pupil dilation

Most eye trackers, applying a pupil plus corneal reflection method to compute the direction of gaze, automatically record the estimated size of the pupil. Pupil diameter is commonly

measured as the diameter (assuming equal height and width of the pupil) or surface area (the number of pixels in the image estimated to be part of the pupil). Analysis of these pupil measurements have suggested that pupil dilation can be used as an indicator of workload, or of emotional responses (Beatty, 1982; Beatty & Lucero-Wagoner, 2000; Bradley et al. 2008; Partala & Surakka, 2003), although pupil dilation is also strongly related to low level visual factors of the stimulus, such as its luminance (the light reflex). For these reasons, pupil dilation data need to be interpreted with care (Bradley et al., 2008). Recently, it has been suggested that pupil dilation is not only linked to the present fixation, but also to the future fixation, opening up the possibility of using pupil dilation to measure covert attention, or the planning of an eye movement (Mathot, Siebold, Donk, & Vitu, 2015). Because pupil dilation data is noisy and depend on the settings of the eye tracker, mean differences with respect to a reference interval are normally analyzed, although some have focused on peak differences, but then averaging across trials (Beatty & Lucero-Wagoner, 2000).

3.4.10. Heatmaps

As a first step to understand where observers focus as a group, heatmaps may be useful (see an example in Figure 3.1), although it has been stressed that they need to be considered with care (Bojko, 2009). Heatmaps use colours to indicate which areas of a scene or image were most often fixated, and thereby provide a visual tool to understand group differences. Statistical methods to compare heatmaps have been developed (Caldara & Mielle, 2011), but these are not without criticism (McManus, 2013). Heatmaps have the advantage that they do not rely on the use of AOI and therefore provide a summary of the data that is not linked to the interpretation of the researcher about what are important areas of the scene. They also provide a useful method of visualizing where observers look when defining areas of interest

is time-consuming, such as for moving images (as in the Dynamic Images and Eye Movements (DIEM) project (The University of Edinburgh, 2009)). Heatmaps do not take into consideration in which order participants viewed a scene or image, although they can incorporate how often an area was visited (Coco, 2009; Wu, Anderson, Bischof & Kingstone, 2014).

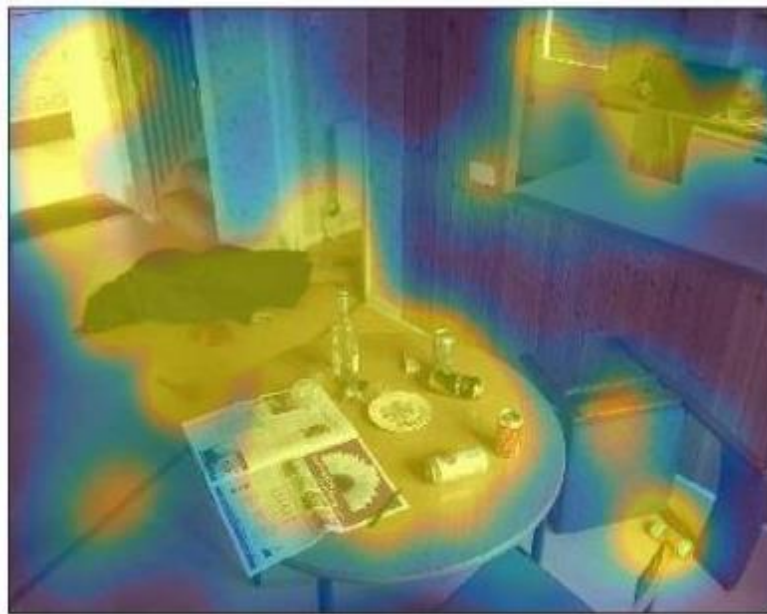


Figure 3.1. An example of an Heatmap from the study, generated using a custom-built MATLAB script that pools 2D Gaussians around individual fixations (irrespective of their duration). Yellow areas show frequently fixated areas. No colour was used when the area was not fixated.

3.4.11. Recurrence Quantification Analysis (RQA)

As mentioned earlier, methods have been developed to deal with both the temporal and spatial aspects of fixation sequences. An important new development in this context is RQA

(Anderson, Bischof, Laidlaw, Risko & Kingstone, 2013; Vaidyanathan, Pelz, Alm, Pengcheng & Haake, 2014). This method has the advantage over scanpath methods that it automatically quantifies the overlap between two patterns that are not necessarily aligned in time (i.e., it can deal with delays between patterns, which is particularly useful when analyzing social interactions, where one viewer can lead the other viewer at a certain interval). RQA produces a range of measures, such as Recurrence, Determinism and, Laminarity. Recurrence measures how often a fixation sequence occurs. Determinism measures the variability of these sequences. Laminarity measures how often a specific location is refixated. (Anderson et al., 2013; Vaidyanathan et al., 2014; Wu et al., 2014). From these measures, Recurrence provides information about the global characteristics of eye movement behaviour and the other two measures, Determinism and Laminarity, provide a measure of the temporal aspects of the eye movement sequences (Vaidyanathan et al., 2014). Recurrence analysis can be used to compare a sequence with itself (RQA) or to compare to sequences with each other (cross recurrence quantification analysis; CRQA).

RQA works by providing measures of when and how often the eye revisits certain areas of the scene. For this to work, one needs to define what these areas are. Two broad approaches to labelling areas have been put forward (Anderson et al., 2013). On the one hand, there are the Fixed-Grid Methods. As the name suggests, Fixed-Grid Methods draws grids onto the scene and provides statistics on eye movements with respect to these grids. The main issue with fixed grid methods is to choose the size of the grid. Often objects end up in multiple boxes in the grid, meaning that the gaze revisits an object, they may revisit different grid boxes. Fixed-Grid Methods, however, have the advantage that the areas are all of the same size, and therefore area size is not influencing the measures. A different approach would therefore be to define areas of interest around objects, but this would have the consequence

that areas have different sizes. The second class of methods, namely the Fixation Distance Methods, defines revisits of areas on the basis of the distance of previous fixations to that area. Typically, a distance of 2° of visual angle is used to define refixations of an area, equalling the estimated foveal vision size (Rayner, 1998), which will also be applied in the present work.

3.5. Eye Tracking Techniques

In order to measure where an observer is looking an eye tracker is used. Over the years, many eye tracking systems have been developed, which can be classified into a range of types, depending of the signals that they use to estimate where the observer is looking. In the very old systems, lenses were attached to the eye (Yarbus, 1967), and when the lens moved, this was an indication of an eye movement (Figure 3.2). These systems were highly invasive, and could therefore be used in only a small number of participants (often the researchers themselves). In the overview here, I will not go any further into these systems.

3.5.1. Electro-OculoGraphy (EOG)

EOG is one of the older techniques, but is still in use, possibly to its low cost, its high sampling rate, and because it allows for the recording of eye movements while the eye is closed (useful in sleep studies). Electrodes are attached (glued) to the muscles that drive the eye movements. Because of the electrical signals involved in driving the eyes, EOG can determine when a muscle is active and how active it is. With EOG, recorded eye movements are often restricted to horizontal movements only, because, in order to record both dimensions, electrodes would need to be placed left, right, above and below the eye, and the

wires may interfere with vision. EOG is often used in EEG experiments, simply to detect eye movements that could interfere with signals measured from the scalp. EOG is not very invasive, but setting up a participant takes some time. Electrodes used are either single use, or have a limited life-time, so there is a maintenance cost associated with this type of eye tracking. EOG is useful to detect when participants make an eye movement, but to detect where exactly observers are looking, other techniques are more useful.

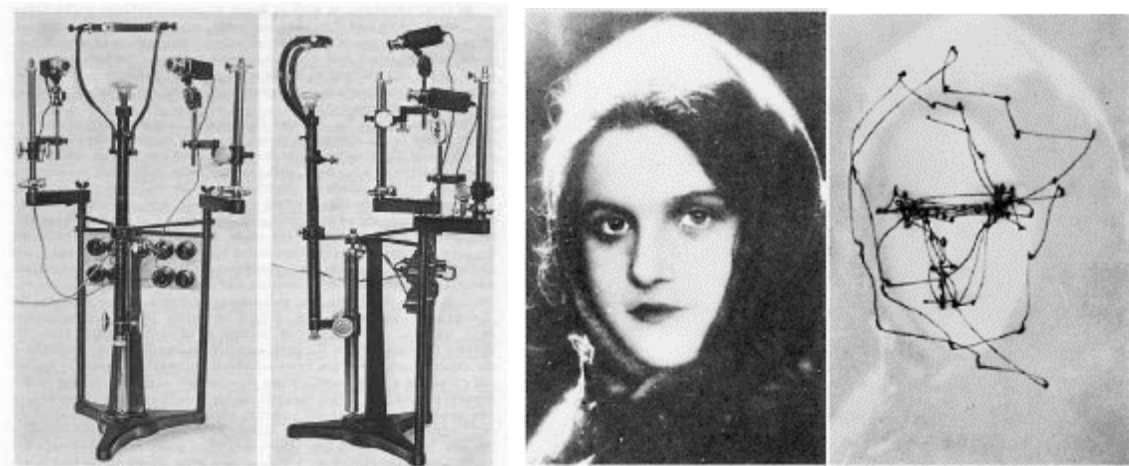


Figure 3.2. An example of an earlier eye movement recording device and recordings (taken from Yarbus, 1967, Figure 21 and 114), with one of the earlier, more intrusive eye trackers.

3.5.2. Scleral Coil

This older technique is still in use today, possibly due to its accuracy (often called the ‘gold standard’ in eye tracking; Collewijn, Van der Mark, & Jansen, 1975; De Bie, 1985), and because until recently, it provided one of the only reliable ways of estimating where observers are looking when they freely move their body (this application has now been taken over by mobile eye tracking). Other reasons for using search coils are to measure torsional eye movements (Ferman, Collewijn, Jansen, & Van den Berg, 1987), e.g., due to movements of

the body (Kaptein & Van Gisbergen, 2006), and in research with primates, where the lens can be implanted in the primate's eye during surgery. To measure eye movements from a search coil, the coil (and observer) need to be placed in a magnetic field. Following Faraday's law of induction, rotation of the coil will induce a small current in the coil, which can be measured. The method is rather invasive. In order to insert the coil into the observer's eye, a local anaesthetic needs to be applied. The coil is often irritating the eye (Murphy, Duncan, Glennie, & Knox, 2001), meaning that testing times have to be relatively short. The wires attached to the coils are very thin and break regularly, and replacing these wires comes at a considerable maintenance cost, not observed with newer systems. Inserting the coil into observers' eyes requires training (Dell'osso & Daroff, 1999), and the analysis of the signals, particularly if participants are allowed to walk around freely, is complex. As a result of these reasons such as the development of more affordable and often similarly reliable methods (Kimmel, Mammo, & Newsome, 2012; Van der Geest & Frens, 2002), and the need for a special magnetic cage to be built inside the experimental room, scleral coil techniques can be used less and less.

3.5.3. Dual Purkinje Image (DPI) eye trackers

DPI eye trackers (Crane & Steele, 1985) are one of the eye trackers in the class of infra-red oculography methods. They make use of reflections of infrared light signals sent into the observer's eye, most often using infrared emitting diodes (IREDs). Four of those reflections can be observed from the human eye, known as Purkinje images. The Purkinje image (P1) arises from the surface of the cornea, the second image (P2) is a reflection from the inner surface of the cornea, the third image (P3) is a reflection from the outer surface of the lens, and the fourth and final image (P4) arises from the inner surface of the lens. From these

signals, the DPI eye tracker makes use of the 1st and the 4th reflections (other trackers often use the 1st reflection in combination with the estimated pupil centre, more about this later). Both reflections move in similar distances during translation of the eye (e.g., due to head movements), but in different proportions during rotations of the eye (i.e., due to eye movements). The vector (directed distance) between the two reflections can therefore be used to estimate the direction of gaze. Despite these properties, DPI eye trackers are often used in combination with a bite-bar, to avoid any head movements in the participant. This makes taking part in an eye tracking experiment with a DPI eye tracker relatively uncomfortable. DPI eye trackers were in use until recently, because of the accurate measurement of eye movements, in particular in the context of measuring fixational eye movements (Collewijn, & Kowler, 2008). However, maintenance of the system is complicated, and therefore associated with a relatively high cost, which is the likely cause for researcher to have abandoned the systems in favour of more recent pupil-corneal reflection systems (more about these below).

3.5.4. Limbus eye trackers

Another class of eye trackers that relies on reflections of infrared light, are the limbus eye trackers. The limbus is the edge where the sclera meets the iris of the eye, and can be detected relatively easily from a video image. The technique requires for the eye to be kept still, and the method suffers from occlusion from the eye-lids, possibly allowing for accurate horizontal eye tracking only (Morimoto & Mimica, 2005). While the limbus is more easily detected than the centre of the pupil, more eye trackers make use of pupil centre signals to estimate the direction of gaze.

3.5.5. Dark and light pupil methods

Most eye trackers now in use, make use of an estimate of the centre of the pupil and the first Purkinje image (P1), also known as the corneal reflection. The typical setup of such eye trackers involve an infrared light source and an infrared sensitive camera. The images of the camera are analyzed for the location of the centre of the pupil and the location of the corneal reflection, and the vector between the two, which is relatively independent of translations, but sensitive to rotations, is used to estimate the direction of gaze. Dark and light pupil methods differ in the direction of the infrared light source. In the light pupil method, the infra-red light source is placed close to the optical axis of the camera, resulting in a bright reflection of the pupil in the camera's image. Dark pupil places the light source further away from the camera axis, and consequently the image of the pupil onto the camera image appears as dark (Morimoto, & Mimica, 2005). Because human do not perceive infrared light, the light source is not interfering with vision. As with other methods, a calibration method is used, involving participants to fixate a series of fixation targets presented ahead of them. This calibration method is used to build a model of how the vector between the corneal reflection and the pupil centre depends on the viewing location, which is then used to predict for any given vector, where (horizontal and vertical coordinates on the screen), the observer is looking (Morimoto & Mimica, 2005; Duchowski, 2007). In addition to the pupil centre, these types of eye trackers often also record the size of the pupil, either in surface area, or diameter of the pupil (Wang, 2011). Eye trackers in this class are relatively expensive to purchase (however, new developments are driving down the price, such as for the EyeTribe system), but they have virtually no maintenance costs, are relatively easy to use, and are not (very) intrusive, explaining their popularity.

3.5.6. Mobile eye trackers

While many of the dark pupil systems above allow head movements to some extent (e.g., Eyelink 1000 Remote, Tobii T120, SMI Red), they do require for stimuli to be presented on a computer screen, limiting the external validity of the findings, as performing certain tasks on a computer screen may not equate with performing the same tasks in the real world. Until recently, mobile eye tracking, in which the observer is able to freely move around, was limited, because technology to do so was often rather complicated to use (as in the search coil systems), relied on custom-built systems (e.g. Land et al., 1999), or was fairly uncomfortable for participants to wear (as for the scene camera Eyelink II system).

Mobile eye trackers, however, have now become commercially available and are often relatively easy to use. They make use of the same principle as the dark pupil systems described above, with an infrared source and an infrared camera, combined with a scene camera recording an image from the point of view of the observer. The infrared source and the camera are often built into the frame of a pair of glasses, with both elements positioned below the observer's eye. Calibration of the systems involves fixating one (Tobii glasses) or multiple (e.g., SMI glasses, Positive Science eye tracker) points in the real world. Mobile eye tracking systems have the advantage that participants are freely able to move around (within the limits of the cables attached to the system), and interact with a 3D world.

A major disadvantage of mobile eye tracking is the labour-intensive nature of the data analysis. Since each participant generates their own "background" images, which change from frame to frame, areas of interest analyses for mobile eye tracking methods take extreme amounts of time. Often software is used for so-called "semantic tagging" in which periods in

which the eye remains relatively still (note that fixation detection is becoming more complicated as well, due to VOR, and OKN signals) are tagged by a human coder as belonging to certain objects in the scene (this can be done if the scene involves a room, and becomes more difficult, for example, in driving studies, where the scene involves a large stretch of road). Further disadvantages are that participants are often aware that their eye movements are being recorded, and it has been suggested that this influences viewing behaviour, particularly in social situations. The frame onto which the camera and light source are mounted often limit the field of view of participants. Mobile eye trackers often have limited spatial and temporal resolution compared to the desk based systems (e.g., 30Hz or 60Hz sampling for mobile systems compared to 500Hz to 1000Hz sampling for desk mounted systems). The lower sampling rates limits the size of differences between groups or conditions that can be detected (Jacop & Karn, 2003).

3.5.7. Combining methods

In order to take advantage of the strengths of the different types of systems (desk mounted and mobile eye tracking), studies could apply both techniques in different sections of the experiment, or across experiments and compare the results. This thesis will make use of this approach, by using a desk mounted system in Experiments 1, 2 and 4 and a mobile eye tracker in Experiment 3. If the both types of tasks (screen-based, real-world) tap into similar aspects of expertise in CSI, I expect results to be similar across the different experiments.

3.6. Equipment in current study

This dissertation will make of two systems, the Eyelink 1000 system (desk mounted system) and the Positive Science Eye Tracker (mobile eye tracker). In the following, a more detailed description of both systems will be given.

3.6.1. The Eyelink 1000 Desk Mounted Eye Tracking Device

The Eyelink 1000 desk mounted system is produced by SR Research located in Canada. The system employs a monocular camera and an infrared light source. The camera and light source are both directed to the participant's head, which can (for more accurate recordings), but does not have to be, stabilized by means of a chin rest (see Figure 3.3). A calibration procedure needs to be conducted at the beginning of the study, requiring participants to fixate a series (typically 9) of fixation targets. Using the measurements of the pupil and the corneal reflection during fixation, an internal model is built that links the participants' eye parameters to the screen coordinates. Using a state-of-the-art gaming computer, the system achieves a sampling rate of 1000Hz (stabilized head) or 500 Hz (head movements allowed) with a spatial accuracy of around 0.5° accuracy. The raw eye movement data are automatically parsed into fixations, saccades and blinks. The approach that the Eyelink system takes for this is to find the sections of the trace in which the estimated velocity and acceleration of the eye exceeds predefined thresholds, defining the start of saccades. A similar procedure is used to find the end of saccades (but with deceleration rather than acceleration). Subsequently, blinks are detected as periods in which the pupil is occluded. The remainder of the signal is coded as fixations. Besides recording where on the screen participants look, the system also collects data about the pupil size (either stored as a diameter or a surface area), which can be

used when inspecting the data for effects of emotional content, and of effort related to the task.



Figure 3.3. An Image of an Eyelink 1000 desk mounted eye tracking system with the chin rest. On the left is where the observer sits, with their head in the chin and head rest, looking at the computer screen, showing the stimulus (a crime scene here). In the middle are the two computers, one used for stimulus presentation (left computer), and the other for recording and parsing the eye movement data (right computer). The experimenter sits on the right and has a view of the image shown with a cursor superimposed indicating the estimated viewing direction of the observer. A view of the image of the eye is provided as well, providing a monitor to determine whether settings are still optimal for accurate eye tracking data.

While the Eyelink 1000 allows for sampling at a very high sampling frequency (1000Hz) with high spatial accuracy, the system has the drawback that a chin rest is needed to achieve this high accuracy, which may restrict the ecological validity of the viewing situation. A “remote” option is provided (at an additional financial cost), which tracks the participant’s

head by means of a sticker attached to the person's forehead, but sampling in this mode is at a lower rate (500Hz) and often less accurate. When having younger or older participants, the remote option, however, can be preferred, as when testing participants, who are often less comfortable with using a chin rest.

While SR Research provides some information about how the calibration and saccade detection algorithms work, the exact underlying mechanisms for these procedures are not revealed, which is the norm for commercial systems (Duchowski, 2007). The users, therefore, have to rely on these systems, or embark on the enterprise of coding their own saccade detection algorithm.

To set up a study on the Eyelink 1000 system, various pieces of software are available. These include the software that is included with the system, called Experimental Builder, and the others such as Opensesame (Mathôt, Schreij, & Theeuwes, 2012), the Psychophysics Toolbox running under MATLAB (Cornelissen, Peters, & Palmer, 2002) and C++ coding and also a C library. For the studies presented in this dissertation, the native Experimental Builder software was used, as it is relatively easy to use for not too complex experiments. This software package allows for the programming of experiments by a graphical user interface in which users can drag and place components of experiments together to form the experimental structure. Experimental Builder is based on the Python programming language and therefore allows the use of Python code for functions not implemented by default in the software. Before running an experiment, the software allows for the compilation of the code, allowing for faster running of the code, improving the timing of the experiment.

For data analysis, the Eyelink 1000 system provides software called the DataViewer, which allows for the playback of eye movement data as fixations superimposed on the images presented, or in the form of a movie clip showing a cursor going around the screen indicating where participants looked. The Dataviewer allows for the extraction of a range of eye movement parameters (e.g., saccade duration, fixation duration, saccade amplitude) or statistics related to the areas of interest (when defined during programming the experiment in Experimental Builder). Alternatively, the data collected can be converted to an ascii file listing the timestamp, x-coordinate, y-coordinate, and the pupil diameter for each sample, as well as messages indicating the start and end of fixations, blinks, saccades and the onset of screens during the experiment. These ascii data can then be processed further in software such as MATLAB, for example, to conduct a cross-recurrence analysis.

3.6.2. The Positive Science Mobile Eye Tracking Device

For the mobile eye tracking study (Experiment 3), the head-mounted Positive Science Ultraflex Headgear eye tracking system was used (Positive Science, USA). The system comes in the form of a frame of a pair of glasses, to which the scene camera (above the right eye), infrared marker (below the right eye) and the eye tracker camera (below the right eye) are attached. The system is powered by a unit, and recording is done by a laptop attached to a backpack (see Figure 3.4). The system records eye movements from the participant's right eye using the pupil and corneal reflection signals with a sampling rate of 30 Hz. Because the system samples from one eye only, this means that depth information (vergence signals) are not incorporated, and therefore it is recommended to calibrate the system at different viewing distances. The offline procedure to combine the eye signal and the scene camera recordings, together with the calibration, allows for incorporating calibrations at different distances, and

the overlay video can be rendered for the various calibrations. The Positive Science eye tracker allows participants to freely walk around and explore their environment, although the short cable from the head-gear to the back-pack was reported by some of the participants as impairing their movements.



Figure 3.4. Illustration of the Positive Science mobile eye tracker system. The left image shows the head-gear, with the infrared diode and eye camera mounted below the participants' right eye. The image on the right shows the laptop attached to the backpack that serves to collect the eye movements and for controlling the system.

The Positive Science system comes with a software package, called Yabus, which automatically combines the images from the scene camera with the eye recordings (Figure 3.5). A second piece of software, called Gazetag, parses the signal into periods of fixations and saccades and provides a tool for the used to code the resulting fixations as belonging to certain areas of interest. To calibrate the system, participants are typically asked to fixate a series of objects in the scene (different sets at different distances can be used, but objects within the same calibration should best be at the same distance), or to fixate a single point

and to rotate the head while doing so. These calibration points are then used to merge the scene camera video and the eye gaze data. It is recommended that calibration is performed several times during the experiment, as the headgear can shift during navigation, invalidating the current calibration (Evans, Jacobs, Tardino & Pelz, 2012). For the study, I decided to calibrate the system both at the start and end of the task, for each participant.

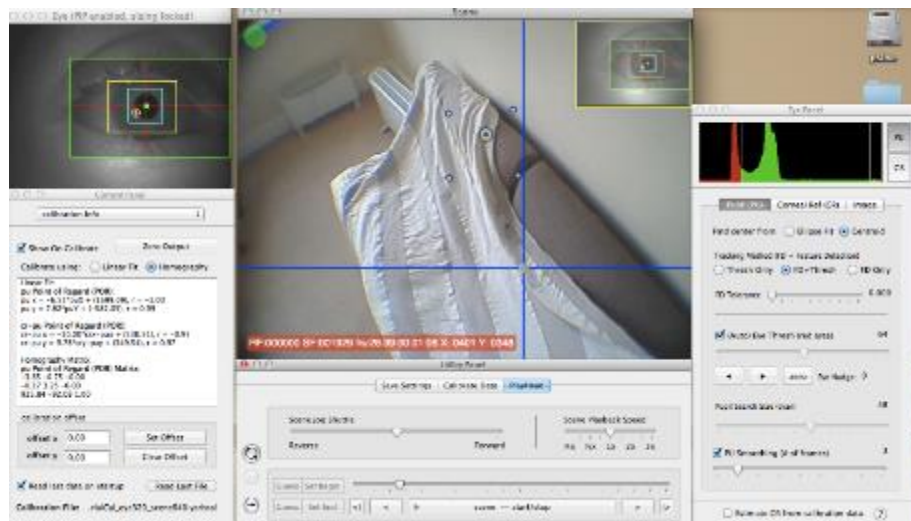


Figure 3.5. Illustration of the Yabus Software. On the left, the signal from the eye camera is shown together with markers to indicate where the centre of the pupil is estimated as well as the corneal reflection. The centre image shows the scene camera view, together with a cursor indicating the estimated gaze location (blue set of lines) and the calibration points (white circular dots). On the right analysis of the signal is shown.

3.7. Non-eye movement methods

To complement the eye movement recordings, participants were also asked to verbally record their trains of thought during the experiment (Hyöna, 2010). This was done by playing the eye movement video of each participant to the participant, and by asking the participant to

comment on their thought processes during exploration of the scenes on the basis of the watched video. Verbal reporting is a commonly used method, and has been suggested as a key method for tracing cognitive reasoning in decision making and problem solving (Ericsson, 2006).

I applied the most common approach, called the “Thinking Aloud” technique (Ericsson & Simon, 1993). While it is possible to ask participants to think aloud during the task, I decided for retrospective thinking aloud, in order not to interfere with conducting the crime scene investigation process. A specific retrospective reporting method was employed, called cued retrospective reporting, where participants report their thoughts while records of the performance are shown to the participant (Van Gog, Paas, Merrienböer & Witte, 2005).

The alternative to retrospective reporting is concurrent reporting. This method places high demands on working memory and may affect performance. As a consequence, concurrent reporting may change the subject's routine and may increase the time required to finish the task (Ericsson & Simon, 1993). Despite these criticisms about the concurrent verbal technique, studies have suggested that the method can give valid inferences about one's cognitive processing (Van Gog, Kester, Nieuvelstein, Giesbers & Paas, 2009).

Retrospective reporting also has some shortcomings and limitations. Retrospective reporting may lead to the fabrication of thoughts that were not in the mind of participants during performance. Furthermore, there may be an inferential bias due to temptation of participants to infer what they have thought (Ericsson, 2006). Another limitation of retrospective reporting is the risk to forget one's thoughts after finishing the task. These risks are particularly high for tasks that take longer and are more complex (Van Gog et al., 2009).

Also, verbal reporting after conducting the crime scene investigation might increase fatigue compared to concurrent reporting (Eger, Ball, Stephens & Dodd, 2007).

A direct comparison of concurrent and retrospective reporting suggests that the two techniques may lead to differences in the information retrieved (Chi, 2006; Van Gog et al., 2009). Van Gog et al. (2009) suggested that these differences may be the consequence of the different role of memory in the two techniques. Concurrent reporting relies on working memory, whereas retrospective reporting technique may rely on LTM (Van Gog et al., 2005). Therefore, concurrent reporting might lead to more action based verbalizations, but retrospective verbalizations might reveal information about strategies used (Van Gog et al., 2005; Van Gog et al., 2009). Another difference between the concurrent and retrospective reporting are the incomplete verbalizations while thinking aloud during the task (Eger et al., 2007). Incomplete and negative responses in concurrent reporting may be the result of the higher cognitive load.

To cope with the disadvantages of both concurrent and retrospective reporting, the cued retrospective verbal technique was employed. In particular, participants were asked to report their thought processes while watching their eye movement recordings. Such cued reporting can help participants to remember thoughts or actions during performance, possibly reducing the fabrication of thoughts and increase the likelihood of reporting thoughts (Van Gog et al., 2009). Eye movement recordings are often used in this context, because they often lead to richer explanations and better process tracing of participants (Duchowski, 2007). Van Gog et al. (2005) showed very similar reporting of strategies and actions in cued reporting and concurrent verbalization. The disadvantage of all verbal reporting techniques is the time required to code the data. This coding can be considered as the most difficult aspect of verbal

reporting techniques (Chi, 2006; Tricket & Trafton, 2007). After transcribing all participants' data, a reliable (which can be interpreted differently by different researchers) coding scheme needs to be constructed (Chi, 2006). These codes need to be linked to the current goal and task of the verbalizer (Tricket & Trafton, 2007). Because coding uses categories, not all verbalizations need to be coded, reducing the overall workload (Chi, 2006). Coding may be susceptible to researcher bias, possibly related to the incomplete verbalizations. Coders therefore need to be aware of misinterpretations and subjectivity. Minimizing the bias effect is of utmost importance (Chi, 2006; Tricket & Trafton, 2007).

3.8. Conclusion

The main approach in this thesis is eye tracking, and in this chapter I have provided an overview of the different eye tracking methods, and the equipment used in this study. Each eye tracking technique has its advantages and disadvantages. In order to best make use of the advantages of the various methods, and to test the robustness of the results under various test conditions, different eye tracking systems and different experimental setups and tasks will be used. Tasks varied from change detection, crime scene investigation to memory tasks, using both static stimuli (photographs) and dynamic situations (using mobile eye tracking). The goal will be to determine how expertise (with crime scene investigation) influences eye movements made during the different tasks. Based on the literature (Chapter 2), expert influences are mostly expected on top-down aspects, whereas bottom-up (stimulus driven) influences are expected to be similar across experts and novices.

Chapter 4. Eye Movements in Experts and Novice Crime Scene Investigators during Viewing of Images of Simulated Crime Scenes

4.1. Introduction

Perception and cognitive reasoning processes are an essential part of the crime scene investigation process, although they have rarely been addressed by empirical research due to the complex nature of crime scene investigation as mentioned earlier. Textbooks on crime scene investigation and forensic science often ignore the role of cognitive processes, and instead focus on the role of technical support.

In the present study, I used eye tracking to gain a better understanding of the perceptual and cognitive processes in CSIs. The goal of the study is to explore how CSIs view and assess crime scenes. This is done by recording the CSIs' eye movements while they perform an assessment of crime scenes depicted in photograph. By comparing the CSIs' eye movement to those of undergraduate forensic science students and controls without CSI experience performing the same task, a better understanding is sought of how expertise influences the visual inspection of a crime scene.

To my knowledge, the present study is the first to apply eye tracking to study the crime scene investigation process. Past studies have predominantly used case based studies or simulated scenarios and have focused on performance to assess expertise effects (e.g., Helsloot & Groenendaal, 2011). Often these studies have used qualitative methods. While informative, such methods often have a subjective component in them in the form of the interpretation of

behaviour and responses by the researchers. The present study aims to complement this research by relying on quantitative measures and the latest eye tracking technology.

Eye tracking has been successfully applied in the assessment of other types of expertise effects. For example, studies have compared eye movements of novice and expert surgeons and found differences in their eye movement patterns (e.g. Ahmidi et al., 2012; Richstone et al., 2010). By training novices to use similar eye movement patterns, acquisition of surgical skill could be sped up (e.g. Wilson et al, 2010). Likewise, eye tracking has been used to study expertise effects in various domains ranging from sports, chess playing, radiology to aviation (Gegenfurtner et al.,2011).

The recent developments in eye tracking technology have contributed to the investigation of skills in different domains. When new technology became available (such as with the commercial introduction of mobile eye trackers), researchers interested in expertise effects became the first to adopt this new technology. Current eye tracking technology falls into two broad categories. On the one hand, there are very precise eye trackers (such as the EyeLink 1000 system, the SMI RED500 system) that allow for eye tracking while participants look at stimuli presented on a computer screen. While these devices allow for participants to move their head while performing the computer based task (e.g., applying the remote option of the EyeLink 1000 system), eye tracking quality is substantially better when head movements are restricted by means of a chin rest. On the other hand, there are the head mounted mobile eye trackers (such as the Tobii Glasses, SMI glasses and the Positive Science system), which allow participants to freely navigate a scene. These systems, however, track people's eyes with lower spatial and temporal precision and resolution. To take advantage of the pros and cons of both types of eye trackers, this thesis will apply both types of technologies in order

to determine whether converging evidence can be obtained demonstrating expertise effects in crime scene investigation.

In the first experiment, static images are used allowing the use of high resolution and precision eye tracking. Participants will be seated in front of a computer monitor and will be asked to view images of crime scenes, in which they click on objects and parts of the scene they think could be evidence for the case. The Eyelink 1000 will be used to collect eye movement data during this process, allowing for the sampling of eye movements at a 1000Hz sampling rate with a documented accuracy of around half a degree of visual angle and a resolution of around 0.05 degrees. Photographs of simulated crime scenes were used, so that the visual aspects of the scenes could be better controlled. Photographs are an essential component of the crime scene investigation process. They are used to document the scene for later evaluation after the investigation area has been cleared for its intended use (Fisher & Fisher, 2012). In this respect, crime scene investigation is different from other domains of forensic sciences, where the use of photographs is less common. Photographs are also much less common in other domains in which expertise effects have been investigated, such as in surgery or aviation (Fitts, Jones & Milton, 2005; Law, Atkins, Kirkpatrick, & Lomax, 2004; Wilson et al., 2010). Crime scene investigation therefore presents as a unique task in which inspecting photographs may carry equal importance in uncovering expertise effects as the actual task of investigating a crime scene. Because a single photograph cannot fully capture a crime scene, two photographs of the same scene were used to allow for views of the scene from different angles. To incorporate the variability in possible scenes, three simulated scenes were used. Ideally a larger number of scenes would be included, but simulated crime scenes are difficult to enact, and therefore the number was limited to three.

4.2. Method

4.2.1. Participants

A total of thirty-six participants were recruited with varying levels of expertise of CSI work. Participants consisted of nine CSIs who regularly assess various crime scenes (experienced group; 5 female), nine undergraduate third year forensic students (intermediate group, 7 female), nine undergraduate first year forensic students (novice group 7 female), and nine participants without any experience or training in CSI (controls; 5 female). While the CSIs' and forensic student's age and gender could not be matched (simply because they were in different stages of their career, which progresses with age), age and gender matching was achieved between the CSIs and the control group.

The mean age for CSIs was 40.8 years. These expert CSIs had an average of 14.4 years of experience in examining crime scenes. The mean age of the intermediate CSIs (third year forensic students) was 22.2 years. They had acquired some understanding of the crime scene investigation, but had no professional experience in crime scene investigations. Novice CSIs (first year forensic students) were on average 26.4 years young and had no training or professional experience in crime scene investigations. Control participants had a mean age of 39.6 years and had no training or experience in crime scene investigations.

Recruitment of the participants started by getting in touch with the Area Forensic Manager of the Police Forces to ask for permission to recruit participants for the study (Invitation Letter in Appendix A, sent to the managers). After obtaining permission to recruit, potential participants were approached with the assistance of the manager, using internal e-mail that

included an information letter. Nominees who were interested in participating the study were scheduled by the Manager to take part on certain dates.

Student participants were recruited by means of an e-mail including an information letter with similar information as the letter sent to the potential CSI participants. The e-mails were sent out by one of the supervisors of the project. To recruit the age matched controls, I specifically targeted potential participants at the university to match the CSIs in age and gender. All testing sessions were conducted during the autumn and winter of 2014.

4.2.2. Ethics Statement

The Ethics committee of the School of Life Sciences of the University of Lincoln approved the study's protocol. Participants all provided consent and provided with written instructions and were verbally debriefed at the end of the study.

4.2.3. Materials

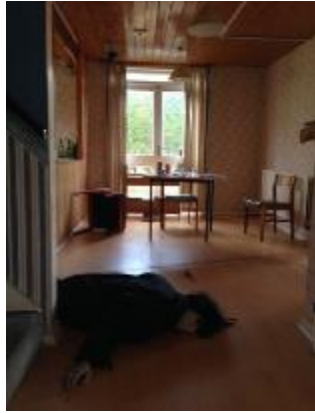
Stimuli consisted of three different simulated crime scene case scenarios. The simulated crime scenes were designed with the help of an experienced CSI. Two of these scenarios were simulated at the Crime Scene House of the University of Lincoln (Riseholme Campus). The remaining simulated crime scene was set up in a spare room in university accommodation at the University of Lincoln. The Crime Scene House has been specially designed to provide realistic learning environment for CSIs and forensic sciences in training. It aims to replicate diverse crime scene scenarios, allowing undergraduate and postgraduate forensic science students to apply previously gained knowledge during lectures into field

work. The two cases simulated in the crime house were located at different locations in the house. Simulated cases were photographed from different angles and two images were selected to be used in the study. Different views of the same crime scene were used to provide views of the same scenes from two angles, which may reveal more details of the scene to the participants than would have been available from a single image and view. The use of multiple images also mimics the crime scene investigation process, where several photographs are typically taken from the same scene, to optimally document the available evidence and layout of the scene. The selected images had spatial resolutions of either 728x1024 pixels (portrait) or 1024x728 pixels (landscape). The images were complemented by a short background story about the case. Detailed information about the simulated scenarios and images is provided in Table 4.1 and the Figure 4.1.

4.2.4. Apparatus

A 22 inch LCD computer monitor (1680x1050 pixel spatial resolution and 43 by 28 degrees size at the viewing distance used) was used to present the stimuli at a 59 Hz refreshment rate. An Eyelink 1000 desk mounted system was used to record eye movements at a sampling rate of 1000 Hz. Participants were seated at a distance of 60 cm from the computer monitor with their head in a head and chin rest (chin rest produced by SR Research, ON, Canada). The scenarios were presented using the Experiment Builder software that came with the Eyelink system under the Windows 7 operating system. Data from the eye tracker were further processed using the DataViewer software supplied with the Eyelink system which produced Excel documents with eye movement statistics. While some of the analyzes were conducted in Excel, others used custom-built MATLAB scripts.

a) Images of Scenario-1



b) Images of Scenario-2



c) Images of Scenario-3

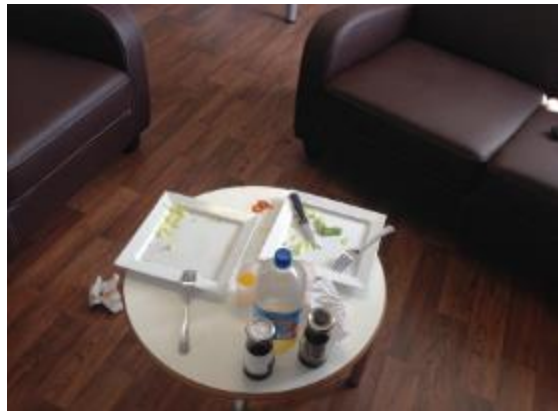


Figure 4.1. a-c) All images used for the three scenarios in the study.

Table 4.1.

The three scenarios which were presented to the participants before the two images for each case.

Scenario-1

A postman reported that, when he was delivering the posts, he found a door open and a woman was lying on the floor, she might have been dead. He didn't touch anything and immediately called the police. Detectives went to the address and learned that she was a widow named Rachelette Disori and living alone. Neighbours mentioned about screams and shouting last night. Scene of Crime requested.

Scenario-2

It has been reported to the police that an employee named Stena Steveson did not come to work in the morning. This was extremely unusual. It was known that she was living alone. Her boss called her several times but there was no reply. He knew that yesterday, when she was at work, her ex-boyfriend called her and she wanted to go home earlier in the afternoon. Officer went to her address but the door was locked. With the help of landlord, the door opened and they found her dead lying on the bed and also there was a bad smell. Scene of Crime requested.

Scenario-3

Nicole Kiddom's mother from Canada called the police and reported that her daughter has not responded her calls for the last two days. She moved to the city two weeks ago and her mother didn't know anyone from there. When they last talked two days ago, she mentioned about a guy, named Carr, who was a student at the same class. Officers went to her address. The door was unlocked. Nobody was at the house but the kitchen was in disorder and probably there were some blood stains in the kitchen. Scene of Crime requested.

4.2.5. Procedure and Design

To facilitate the recruitment of expert CSIs, the experimental setup was temporarily located in the Unit's Headquarters. Experts CSIs were therefore testing in a slightly different setting than the undergraduate forensic students and the control participants, who took part in the Eye Tracking Laboratory at the University of Lincoln. Both settings involved a quiet room with the eye tracker, computers, screen and chin rest positioned on a table. Because participants were new to eye tracker, an introduction to the system was given before the experiment.

The experiment started with an initial practice session, during which the researcher made sure that the participants understood the task requirements. The participants were required to examine the crime scenes in the photographs and to try and identify objects in the scenes relevant to the case. In a second stage, participants were instructed to verbally express their thought processes while watching the playback of their eye movements (the "retrospective think aloud procedure", as discussed in the methods chapter).

To ensure the correct tracking of individuals' eye movements, the eye tracker was calibrated using the default nine point calibration procedure before the display of each scenario. Calibration was considered successful if the nine randomly illuminating calibration points superimposed with eye position and was reported accurate by the Eyelink software. The reported accuracy of eye tracking under such conditions is approximately 0.5° of visual angle on the screen. If the above conditions were not met, the calibration process was repeated until successful. After calibration, participants completed an initial practice session with feedback from the examiner to understand the task in the experiment and to get familiar with the eye

tracker and the process. The practice session involved exactly the same procedure as the real study, containing a simulated scenario which was not used in the main experiment. This practice session also provided the opportunity to ask questions.

After the calibration, scenarios were presented one by one. Each trial started with the short background information presented in the middle of the screen in white font a black background. After participants pressed a key on the keyboard to indicate that they read and understood the background information, the first image. Participants again pressed a key to indicate they were ready to move on to the next image. No time limit was given for reading the case text, or viewing the images. All three scenarios were shown in the same order to all participants. Between the display of the two images, a drift correction was performed, involving participants fixating a centrally presented fixation target, after which the experimenter pressed a key to confirm fixation. This procedure was included to reduce the influence of drift on the eye movement recordings. While participants viewed the images, they used a mouse to click on object/s or place/s possibly containing evidence for the case. These mouse clicks were stored in a separate file in the form of x- and y- coordinates on the screen in the order in which they occurred.

After completing each one of the three scenarios, participants performed a retrospective 'think-aloud' protocol. This protocol involved participants watching their eye movements superimposed on images on the computer screen and to verbalize their thought processes. Because the eye movement recordings were too fast to verbally express one's thoughts in real-time the video clip was played back at reduced speed ($\frac{3}{4}$ of the original speed was chosen), in agreement by the suggestion from Hyrskykari et al. (2008). The recordings were

shown in Dataviewer and a recorder (Smart Phone) was used to record the verbal report, which was analyzed offline. In the protocol, participants focused on two questions:

"Question 1. While you are watching your eye movement recording on the computer screen, provide details about what you were thinking about the picture. Explain your thought process.
Question 2. What is your opinion/s about what happened in this scenario?"

Participants were also provided with information about their mouse clicks and were asked to clarify their choices. After finishing all three scenarios, participants completed a short demographic questionnaire and were thanked for their participation. The experiment took around one hour to complete for each participant.

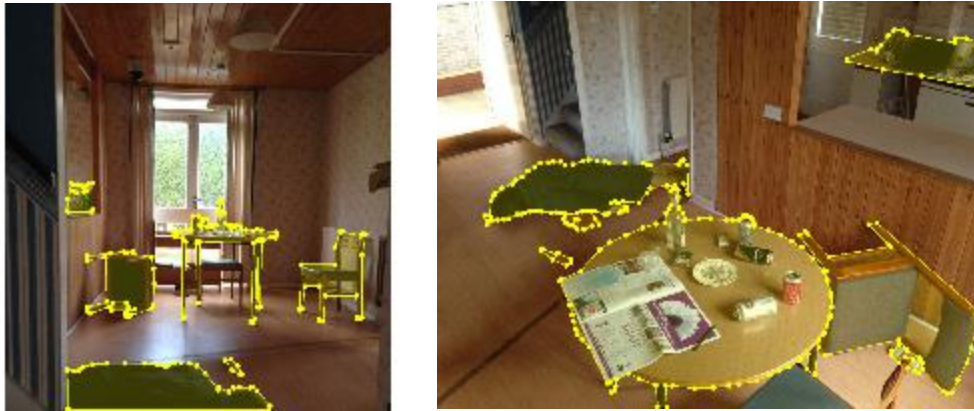
4.2.6. Data analysis

Fixations, saccades, and blinks were extracted automatically from the raw horizontal and vertical eye gaze traces using the in-built parser of the Eyelink system using the default settings. Mean fixation durations, mean saccade amplitudes, number of fixations and the total time duration were extracted from these data for each image and participant using the Dataviewer software. For a fixation to be included into the analysis, it was required it to last at least 100 ms. This minimum duration was adapted from the suggestion by Salvucci and Goldberg (2000), limiting included fixations to those that could lead to meaningful extraction of information.

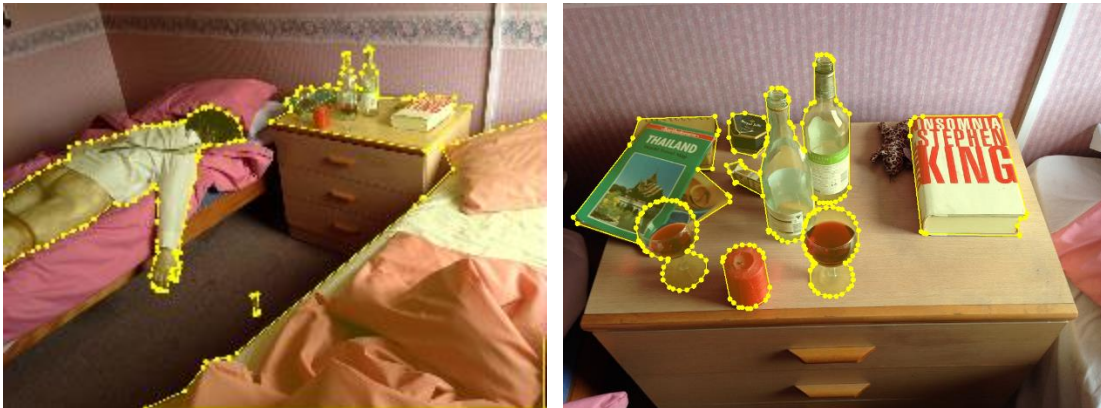
More specific analysis of the eye movement data involved an AOIs analysis. Such an analysis provides more detailed information about the type of objects fixated by the different

participant groups than an analysis of general eye movement parameters can do. The AOIs for the different images are illustrated in Figure 4.2.

a) AOIs in Scenario-1



b) AOIs in Scenario-2



c) AOIs in Scenario-3



Figure 4.2. a-c) AOIs used for analysis for the three scenarios (dashed yellow lines indicate these AOIs).

Further analysis of the eye movement data involved RQA to quantify the temporal order of the eye movements. While RQA can only quantify recurrences in a single sequence, CRQA, can compare pairs of sequences (for details see Chapter 3). Because the aim was to make comparisons across (groups of) participants, CRQA was used.

The CRQA yielded three different measures, each quantifying specific aspects of the recurrence plot, which shows at what times a system enters the same state (e.g., when two observers look at the same area of interest). These measures are: Recurrence, Determinism, and Laminarity (Anderson et al., 2013). Recurrence provides a measure of repeated fixation sequences; Determinism measures how consistent fixation sequences are while Laminarity expresses the refixation characteristics in specific locations in the recurrence plot.

There are various ways to define the spatial locations used for CRQA (Anderson et al., 2013). For example, the analysis can focus on how often participants fixate the same object, or how often they fixate the same area of the screen (areas may or may not include objects). To allow for different size areas, an area-based, fixed grid approach was not able to be used (see the example in Figure 4.3). To take into account the different sizes of possible objects and areas of interest, a fixation distance method was used. In the fixation distance method, fixations were considered as recurrent if the distance between the fixations was within a predefined distance. A 2° of visual angle (foveal vision as defined in Rayner, 1998) distance was used, which equals to 75 pixels as illustrated in Figure 4.4.



Figure 4.3. Illustration of the possible issue when using a fixed Grid Method. The left image shows a 4x4 grid, while the right image shows a 8x6 grid. In neither type of grid, the book occupies a single square, whereas the body spans several squares for both grids. Using such grids would not reveal refixations of individual objects.



Figure 4.4. Illustration of the fixation distance method. Each image shows the analysis of the scanpath of an individual participant (two different participants). Each circle represents one area. Recurrence to that area is defined as a fixation within the 2° of visual angle circle surrounding the center of the area.

Because CRQA can only compare two participants at a time, repeated comparisons between pairs of participants within and between groups were conducted and the recurrence measures

pooled. This yielded average recurrence measures between groups, as well within groups (providing a measure of variability within a group). It was hypothesized that experts would quickly recognize the important objects and would less often refixate these objects, and therefore higher recurrence within experts than between experts and novices was expected. Moreover, low levels of recurrence within experts themselves were expected (less variability). The measures were all computed using the toolbox by Anderson et al. (2013).

To analyze the data for statistical significance, mixed factor ANOVAs were computed, using 'Experience' as a between subjects factor and 'Scenario' and 'Image' as within subjects factors. Because distributions of fixation durations and saccade amplitudes tend to be highly skewed with most observations found at lower values, relatively few large values can have a significant impact on the average data. To avoid such impact, large values were removed from the data by taking out observations more than two standard deviations from the mean for each participant, before computing the average data (e.g., Buckee, Kneller & Peakall, 2001). Subsequently, the data were inspected for violations of normality assumption by the examination of standardized residuals, and by means of Shapiro-Wilk tests. In the case of violations of the normality and the homogeneity of variances assumptions, square root or arcsin transformations were applied to the data before computing the relevant statistics. Furthermore, Huyn-Feldt corrections were applied to the degrees of freedom of each statistical test when the sphericity assumption was violated. The average data plotted, was not transformed, as this would make the data plots difficult to interpret.

Before results were computed, eye movement data of the individual participants were visually inspected for abnormalities. Because some participants showed significant drift in their eye movement data, a manual drift correction was applied to these participants' data

using the corresponding procedure in Dataviewer. For one first year participant a substantial part of the eye movement data was recorded as missing data. Their data were therefore excluded from the final analysis. Also, mouse clicks were analyzed separately. If participants clicked on the same object across the two images, the object was coded only once.

4.2.7. Results

4.2.7.1. Eye Movement Characteristics

Mean Fixation Durations

Mean fixation durations are shown in Figure 4.5a, separately for each image and experience group. For one of the conditions, the homogeneity of variance assumption was violated (Levene's test), and this violation remained after transformation. Therefore, the results should be interpreted with more care.

A mixed factors ANOVA revealed no main effect of experience, despite a slight trend towards longer fixation durations in expert CSIs ($F(3,31)=0.35$, $p=0.79$, $\eta^2 = 0.03$). Fixation durations showed a main effect of scenario ($F(2,62)=281.98$, $p<0.001$, $\eta^2 = 0.9$) and the specific image within each scenario ($F(1,31)=278.82$, $p<0.001$, $\eta^2 = 0.9$). In the absence of an interaction effect between experience and scenario ($F(6,62)=1.77$, $p=0.12$, $\eta^2 = 0.15$), experience and image ($F(3,31)=1.48$, $p=0.24$, $\eta^2 = 0.13$) or experience and scenario and image ($F(6,62)=1.17$, $p=0.34$, $\eta^2 = 0.1$), it was concluded that fixation duration differences between the four groups were not affected by the scenario and image used.

Saccadic Amplitudes

After taking out saccade amplitudes more than two standard deviations from the mean, one participant (an expert) for whom the average saccade amplitude was beyond two times the standard deviation of the mean across participants. To avoid this participant biasing the results, the mean of this participant was replaced by that of participant with the next highest average. Figure 4.5b plots the average saccade amplitudes across images, scenarios and participant groups. Larger saccade amplitudes can be expected for a more global scanning strategy (Tatler & Vincent, 2008) and increased situational awareness in experts may lead to large saccade amplitudes overall.

Analysis of the data showed a main effect of experience ($F(3,31)=3.0$, $p=0.046$, $\eta^2 = 0.23$). Interestingly, the larger saccades were from the third year students, rather than the CSIs. Scenarios led to different saccade amplitudes ($F(1.87,57.85)=14.98$, $p<0.001$, $\eta^2 = 0.33$; $\epsilon=0.93$), possibly because objects of interest were differentially distributed across scenes. Likewise, saccade amplitudes varied between the different images ($F(1,31)=11.2$, $p=0.002$, $\eta^2 = 0.27$). No interaction was found between the different participant groups and the scenarios ($F(5.6,57.85)=1.43$, $p=0.22$, $\eta^2 = 0.12$), the participant groups and images ($F(3,31)=0.23$, $p=0.88$, $\eta^2 = 0.02$), or the participant groups and scenario and image ($F(6,62)=1.06$, $p=0.4$, $\eta^2 = 0.09$).

Number of Fixations

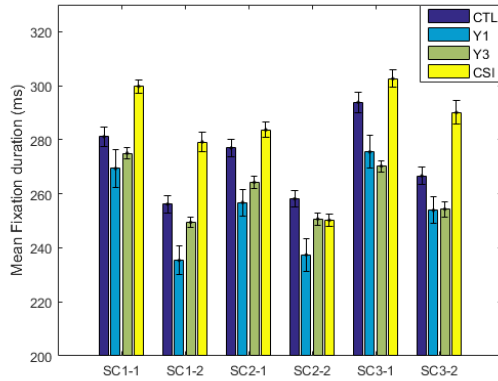
Because the viewing time was not limited, numbers of fixations are not directly related to the fixation durations, and instead rather provide a measure of the intensity with which the image was inspected. Average numbers of fixations are plotted in Figure 4.5c. Statistical analysis of these data did not reveal an effect of Experience, although students showed a trend towards

lower numbers of fixations. Possibly because of the age matching, the number of fixations were similar for CSIs and controls ($F(3,31)=0.28$, $p=0.84$, $\eta^2 = 0.03$). Scenario significantly influenced the number of fixations ($F(2,62)=63.57$, $p<0.001$, $\eta^2 = 0.67$), as did the images ($F(1,31)=55.48$, $p<0.001$, $\eta^2 = 0.64$), without an interaction between experience and scenario ($F(6,62)=0.47$, $p=0.83$, $\eta^2 = 0.04$), between experience and image ($F(3,31)=2.0$, $p=0.14$, $\eta^2 = 0.16$), or between experience and scenario and image ($F(6,62)=0.44$, $p=0.85$, $\eta^2 = 0.04$).

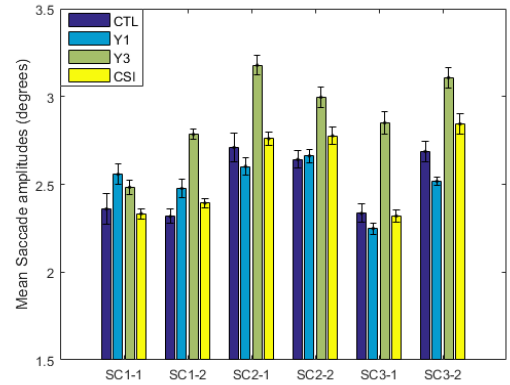
Dwell Time Percentages on AOIs

Although the unsuccessful transformations of the non-normal data, due to the robust nature of violations of normality assumption, the analysis was performed using ANOVA with the untransformed data (Field, 2009). Figure 4.5d plots the time participants spent looking at any of the areas of interest, which were often objects placed by the experimenter, related to the scene. Overall, a large portion of the trials was used to look at these areas of interest. Experience did not have an effect on dwell times towards the different AOIs ($F(3,31)=0.06$, $p=0.98$, $\eta^2 = 0.01$). In contrast, scenario significantly influenced the type of AOIs looked at ($F(2,62)=87.07$, $p<0.001$, $\eta^2 = 0.74$). An effect of images was also found ($F(1,31)=92.19$, $p<0.001$, $\eta^2 = 0.75$). These effects were mediated by a marginally significant interaction effect between experience and scenario ($F(6,62)=2.1$, $p=0.066$, $\eta^2 = 0.17$), suggesting that different scenarios affected the general viewing patterns of different participant groups in specific ways, in the absence of an interaction effect between experience and image ($F(3,31)=0.33$, $p=0.81$, $\eta^2 = 0.03$), or experience and scenario and image ($F(6,62)=1.4$, $p=0.22$, $\eta^2 = 0.12$).

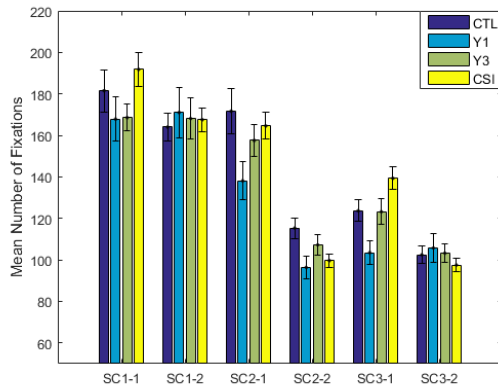
a) Mean Fixation Durations



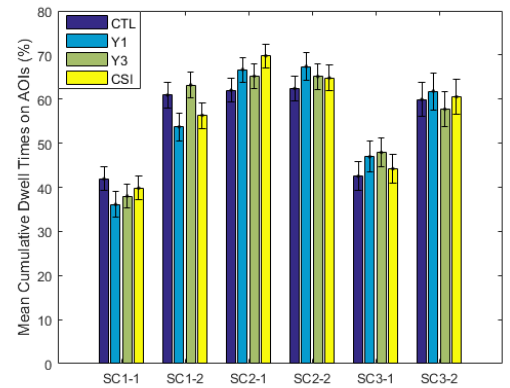
b) Mean saccadic amplitudes



c) Mean Number of Fixations



d) Mean Dwell Times on AOIs



e) Mean Total Duration

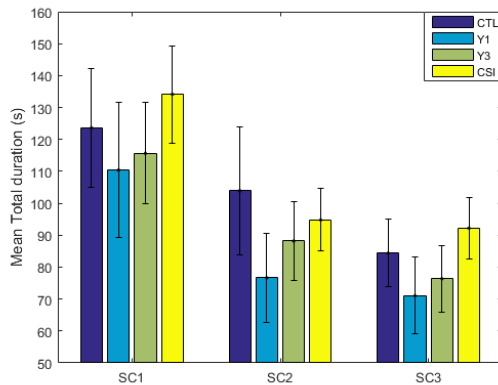


Figure 4.5. a) Mean Fixation Durations (in ms, N=35). b) Mean saccadic amplitudes (in degrees of visual angle). c) Mean Number of Fixations. d) Dwell times towards either of the areas of interest (as a percentage of overall viewing time on the image), plotted separately for each image and scenario (two images per scenario, along the horizontal axis of the plot). e) Mean viewing time (s), plotted separately for each scenario and participant group (different colour; see legend, CTL=Control, Y1=First year students, Y3=Third year students, CSI=Crime Scene Investigators). The error bars show the standard error of the mean.

Total viewing time

Total viewing time computed for each scenario adding the time spent on each image together. Figure 4.5e plots the average viewing time for each scenario per participant groups. No significant effect of expertise was found ($F(3,31)=0.96$, $p=0.43$, $\eta^2 = 0.09$), however scenario did influence viewing time ($F(2,62)=54.54$, $p<0.001$, $\eta^2 = 0.64$). There were no interactions between experience and scenario ($F(6,62)=0.62$, $p=0.71$, $\eta^2 = 0.06$), indicating total time spent on each scenario were not affected by the experience differences.

CRQA of the Eye Movement Sequences

Recurrence, Determinism and Laminarity measures were compared between participant groups, involving the following comparisons: CSIs-First Year Students[CSI-Y1], CSIs-Third Year Students[CSI-Y3], CSIs –Controls[CSI-CTL] and CSIs-CSIs[CSI-CSI]. Due to the severe violations of normality and homogeneity assumptions and the failed transformations of the non-normal data and also the unbalanced group sizes, the results in this section need to be interpreted with care. As the results of the transformed and untransformed data were similar, the statistical analyzes of the untransformed data will be presented here alone. While no non-parametric candidates are available for the mixed ANOVA analyzes, direct comparisons of individual groups could be performed with non-parametric tests (Kruskall-Wallis test), which reduce the influence of violations of normality and homogeneity. An example of a recurrence plot and fixation sequences were shown in Figure 4.6. and the averages in CRQA measures for each scenario per participant groups were plotted in Figure 4.7.

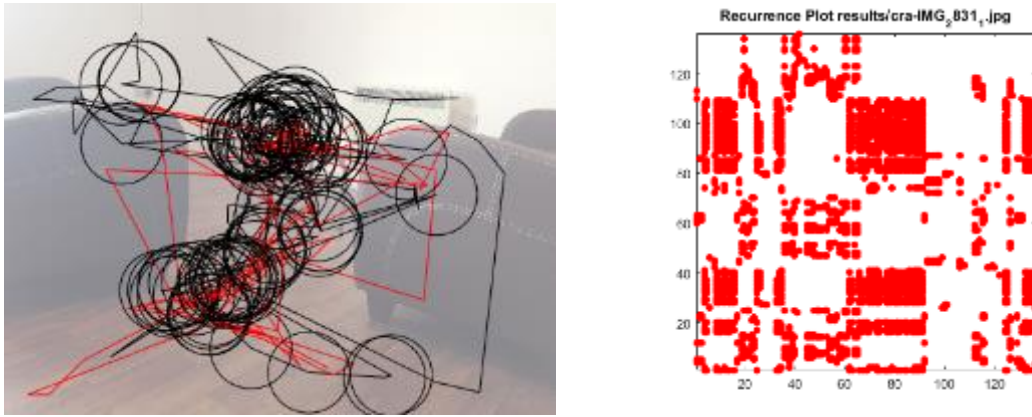


Figure 4.6. Example of fixation sequences using fixation distance method (Left) and recurrence plot (Right) of a CSI vs. CSI comparison. Black and red lines represent different participants in the images. Red points show the temporal fixation sequences in the plots.

The recurrence measure revealed no effect of expertise ($F(3,266)=0.95$, $p=0.42$, $\eta^2 = 0.01$). Recurrence was affected by scenario ($F(1.42,376.72)=192.51$, $p<0.001$, $\eta^2 = 0.42$; $\epsilon=0.71$) and the image used ($F(1,266)=571.19$, $p<0.001$, $\eta^2 = 0.68$). There was an interaction between scenario and experience ($F(4.25,376.72)=7.51$, $p<0.001$, $\eta^2 = 0.08$). However, no interaction was found between experience and image, ($F(3,266)=1.82$, $p=0.14$, $\eta^2 = 0.02$; $\epsilon=0.81$) or experience and scenario and image ($F(4.85,429.94)=1.78$, $p=0.12$, $\eta^2 = 0.02$).

Determinism showed a main effect of experience ($F(3,266)=9.34$, $p<0.001$, $\eta^2 = 0.1$). Games Howell Tests corrected the only significant differences between group comparisons for C-S3 group. (C-S3 vs. C-S1, $p<0.001$; C-S3 vs. C-Cont, $p=0.001$ and C-S3 vs. C-C, $p<0.001$) A further main effect was found of scenario ($F(1.79,477.14)=70.77$, $p<0.001$, $\eta^2 = 0.21$; $\epsilon=0.90$) and of the image used ($F(1,266)=1420.08$, $p<0.001$, $\eta^2 = 0.84$). There was an interaction between scenario and experience ($F(5.38,477.14.)=6.25$, $p<0.001$, $\eta^2 = 0.07$), in the absence of an interaction between experience and image ($F(3,266)=0.54$, $p=0.65$, $\eta^2 =$

0.01). There was a significant interaction between experience and scenario and image ($F(6,532)=2.91$, $p=0.008$, $\eta^2 = 0.03$).

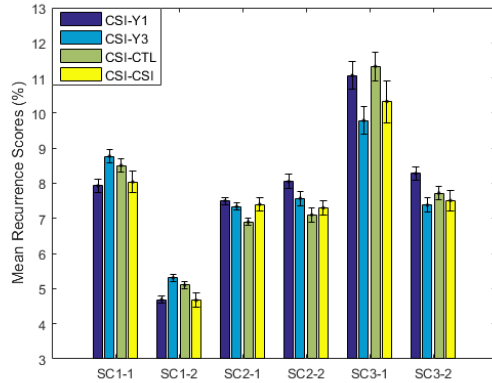
Further analysis for each image in each scenario indicated differences on viewing patterns on Kruskal-Wallis non parametric tests. (SC1; Image-1 $H(3)=10.65$, $p=0.014$, Image-2 $H(3)=7.55$, $p =0.056$; SC2; Image-1 $H(3)=21.77$, $p<0.001$, Image-2 $H(3)=14.02$, $p=0.003$; SC3; Image-1 $H(3)=18.76$, $p<0.001$, Image-2 $H(3)=17.24$, $p=0.001$) Pairwise comparisons showed differences between some of the groups (SC1; Image1, CSI-Y1 vs. CSI-CSI, $p=0.004$; SC2; Image1, CSI-Y1 vs. CSI-Y3, $p=0.043$, CSI-Y3 vs. CSI-CTL, $p=0.003$; Image2, CSI-Y1 vs. CSI-Y3, $p=0.001$; SC3, Image1, CSI-Y1 vs. CSI-Y3, $p=0.001$, CSI-Y3 vs. CSI-CTL, $p=0.002$)

There was no effect of experience on Laminarity scores ($F(3,266)=0.91$, $p=0.44$, $\eta^2 =0.01$). A main effect was found of the scenario ($F(1.87,497.84)=71.03$, $p<0.001$, $\eta^2 = 0.21$; $\epsilon=0.94$) and the image used ($F(1,266)=590.8$, $p<0.001$, $\eta^2 = 0.69$). There was no interaction between scenario and experience ($F(5.62,497.84.)=1.65$, $p=0.14$, $\eta^2 = 0.02$) and experience and image ($F(3,266)=0.394$, $p=0.76$, $\eta^2 = 0.004$). However, there was an interaction between experience and scenario and image ($F(5.74,509.07)=3.06$, $p=0.007$, $\eta^2 = 0.03$; $\epsilon=0.96$) meaning that temporal refixation sequences showed different patterns between groups in each image of the scenarios.

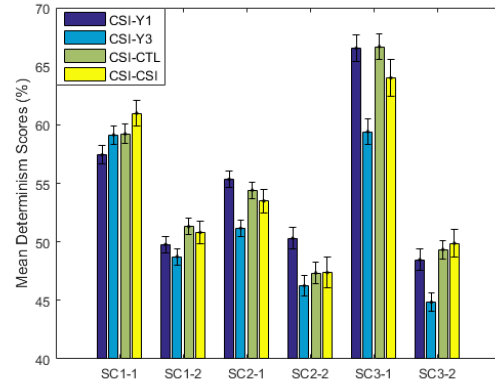
Further analysis for each image in each scenario indicated trends on differences in the all images on eye movement sequences on Kruskal-Wallis non parametric tests (SC1; Image-1 $H(3)=7.9$, $p=0.048$, Image-2 $H(3)=6.36$, $p =0.095$; SC2; Image-1 $H(3)=7.8$, $p=0.05$, Image-

2 $H(3)=8.31$, $p=0.04$; SC3; Image-1 $H(3)=9.82$, $p=0.02$, Image-2 $H(3)=2.64$, $p=0.45$) and however, no significant group differences were found on pairwise comparisons.

a) Recurrence



b) Determinism



c) Laminarity

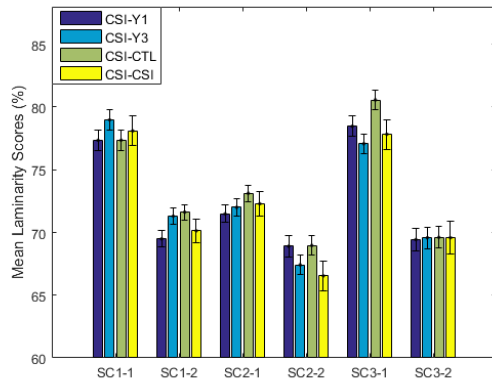


Figure 4.7. Average CRQA measures for each image between group and within group (N=35): Recurrence, Determinism and Laminarity. CSI-CSI: Within group CSI comparisons, CSI-Y1: Between group CSI, first year student comparison, CSI-Y3: Between group CSI, third year student comparison, CSI-CTL: Between group CSI, control group comparison. Error bars show the standard error of the mean across participants

4.2.7.2. Selected Evidence

Mouse clicks were not accurately stored for two participants (one CSI and one third year student), which were removed from the analysis, in addition to the data of the first year participant who was excluded for poor eye tracking data.

The homogeneity of variance assumption was violated (Levene's test) for one of the conditions, which persisted after transformation of the data. Therefore, interpretation of the ANOVA results should be considered with care. Figure 4.8 plots the number of objects selected as evidence across the different participant groups and images. Scenario influenced the number of objects selected as evidence ($F(1.71,49.71)=50.27$, $p<0.001$, $\eta^2 = 0.63$; $\epsilon=0.86$) and an interaction was found between experience and scenario ($F(5.14,49.71)=2.76$, $p=0.027$, $\eta^2 = 0.22$). CSIs showed a trend towards selecting more evidence than the other participant groups, but this trend was not statistically significant ($F(3,29)=2.37$, $p=0.92$, $\eta^2 = 0.2$).

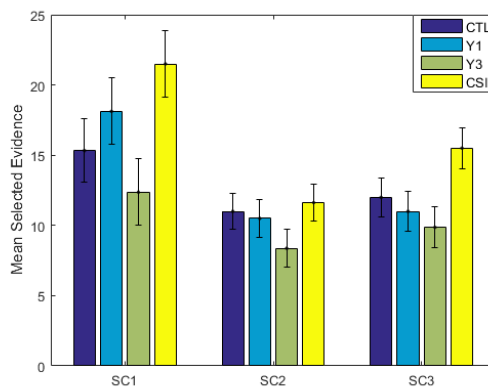


Figure 4.8. Number of objects selected as evidence for each scenario (horizontal axis) and participant group (different colours, see legend, CTL=Control, Y1=First year students, Y3=Third year students, CSI=Crime Scene Investigators, N=33). Error bars show the standard error of the mean across participants.

Further one-way ANOVA analysis also showed that, CSIs showed a trend towards selecting more evidence than the other groups, specially for Scenario-1 ($F(3,29)=2.71$, $p=0.063$, $\eta^2=0.22$) and Scenario-3 ($F(3,29)=2.83$, $p=0.056$, $\eta^2=0.23$). Games Howell Post hoc tests showed a difference only between CSIs and third year students (Scenario-1, $p=0.005$; Scenario-3, $p=0.012$). No such group differences were found for Scenario-2 ($F(3,29)=1.12$, $p=0.36$, $\eta^2=0.1$). Kruskal Wallis non-parametric comparisons showed similar trends with the one way ANOVAs (Scenario-1, $H(3)=7.93$, $p=0.048$; Scenario-2, $H(3)=2.83$, $p=0.42$; Scenario-3, $H(3)=8.22$, $p=0.042$; Pairwise comparisons SC1, CSI vs. Y3 $p=0.023$; SC3, CSI vs. Y3 $p=0.027$).

4.2.7.3. Verbal Protocols

Verbal reports from the retrospective cueing paradigm were analyzed for all but one CSI, where the recording of the report was of such poor sound quality that it was unclear what the participant said. Verbal data were categorized on the basis of the protocol by Trickett and Trafton (2007) and Baber and Butler (2012). The following categories were used:

- a. Reference to Objects in the scene. Knife, glass, cups etc.
- b. Reference to Scene and Scene's features. General parts of the room, door, bed etc.
- c. Reference to Modus Operandi. Perpetrator's or other activities that inferred about the case
- d. Reference to analysis or evidence. Possible explanations for plans for further examinations (e.g. DNA, fingerprints)

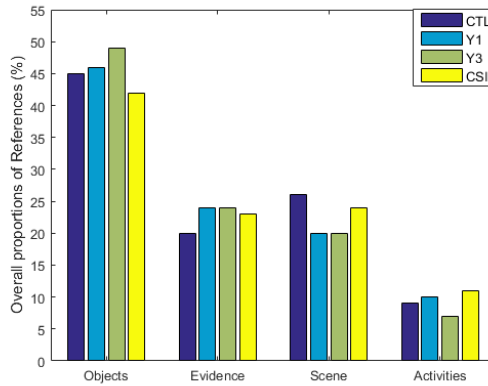


Figure 4.9. Overall proportions of coded retrospective reports in each participant group (CTL=Control, Y1=First year students, Y3=Third year students, CSI=Crime Scene Investigators)

To obtain a measure of the reliability of the coding, a second coder (an experienced CSI with extensive knowledge and experience in qualitative analysis) coded 25% of the transcripts independently. Disagreements between coders were resolved by discussion. The Mezzich's K statistic (Eccleston, Werneke, Armon, Stephenson & MacFaul, 2001) showed a value of $\kappa = 0.64$, indicating substantial agreement between the raters (Landis & Koch, 1977). A coding example can be found in Table 4.2:

(CSI-5 reporting in Scenario-2) “..... Two people have been drinking, there are two bottles of wine. Just DNA swabbing of the glass, fingerprint and the other glass..... Have they been reading that book? The surface is good for at least some sort of fingerprint recovery. And is the cigarette. Again, looking to see if there is an ashtray. Err, when I saw the brochure I thought maybe actually looking and again back together go on a holiday type thing.....”

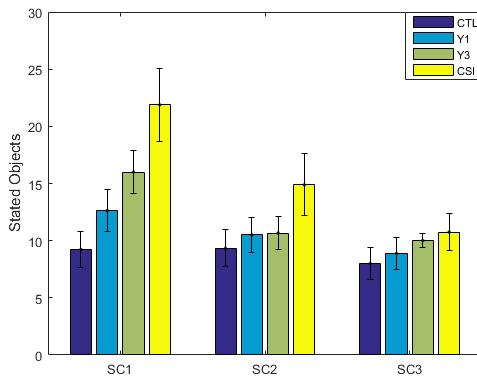
Table 4.2.

An example of coding process.

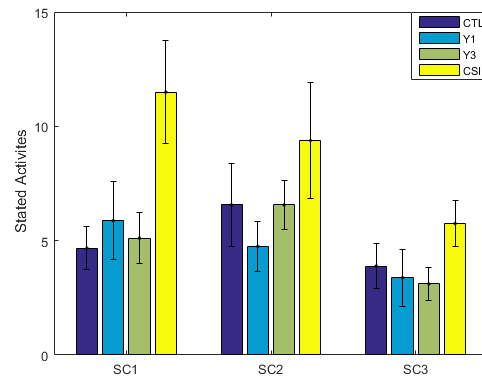
<i>Coding Scheme</i>	
<i>Referring to object</i>	<ol style="list-style-type: none"> 1. There are two bottles of wine 2. glass ... and the other glass. 3. Book surface 4. the cigarette 5. When I saw the brochure
<i>Referring to evidence</i>	<ol style="list-style-type: none"> 1. Just DNA swabbing... fingerprint 2. good for at least some sort of fingerprint recovery
<i>Referring to Activities</i>	<ol style="list-style-type: none"> 1. Two people have been drinking 2. Have they been reading that book? 3. to see if there is any ashtray 4. maybe actually looking and again back together go on a holiday type thing.
<i>Referring to scene</i>	

Naming frequencies are shown in Figure 4.10. CSIs made more references overall ($F(3,30)=3.63$, $p=0.024$, $\eta^2 = 0.27$). Scenario also influenced the number of references ($F(2,60)=64.87$, $p<0.001$, $\eta^2 = 0.68$) and frequencies of naming also varied across Types of References ($F(3,90)=78.29$, $p<0.001$, $\eta^2 = 0.80$). These main effects are modulated by an interaction between experience and scenario ($F(6,60)=3.57$, $p=0.004$, $\eta^2 = 0.26$) and experience and scenario and type of references ($F(18,180)=1.94$, $p=0.015$, $\eta^2 = 0.16$). There was no interaction effect between experience and type of references ($F(9,90)=1.08$, $p=0.39$, $\eta^2 = 0.1$).

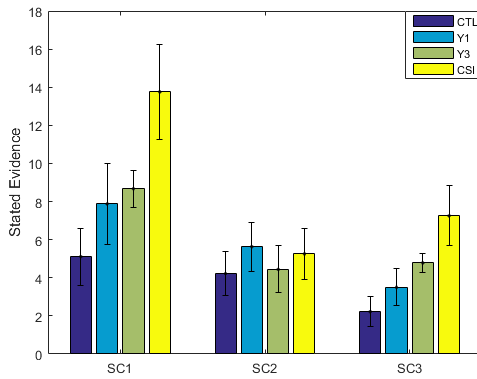
a) Objects stated in Verbal Reports



b) Activities stated in Verbal Reports



c) Evidence stated in Verbal Reports



d) Scene features stated in Verbal Reports

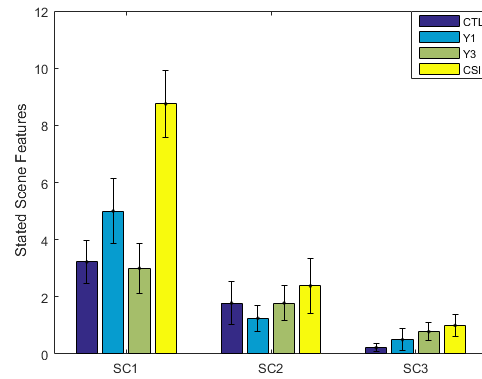


Figure 4.10. a-d) Average references stated in verbal reports for each scenario between groups (CTL=Control, Y1=First year students, Y3=Third year students, CSI=Crime Scene Investigators, N=33).

CSI made more references than the other groups, for Scenario-1 (Objects, $F(3,30)=6.04$, $p=0.02$, $\eta^2 = 0.38$); Evidence, $F(3,30)=3.82$, $p=0.02$, $\eta^2 = 0.28$); Inferences, $F(3,30)=3.55$, $p=0.026$, $\eta^2 = 0.26$); Scene; $F(3,30)=5.32$, $p=0.005$, $\eta^2 = 0.35$), but not for Scenario-2 (Objects, $F(3,30)=1.43$, $p=0.25$, $\eta^2 = 0.12$); Evidence, $F(3,30)=0.49$, $p=0.69$, $\eta^2 = 0.05$); Inferences, $F(3,30)=1.14$, $p=0.35$, $\eta^2 = 0.1$); Scene; $F(3,30)=0.38$, $p=0.77$, $\eta^2 = 0.04$). For Scenario-3 reference to evidence was different across participant groups ($F(3,30)=1.11$, $p=0.36$, $\eta^2 = 0.36$), reflecting a difference between CSIs and controls ($p=0.003$, Tukey HSD

Post Test). However, no other differences were found in any of the comparisons. (Objects, $(F(3,30)=1.43, p=0.25, \eta^2 = 0.1)$; Inferences, $(F(3,30)=1.74, p=0.18, \eta^2 = 0.15)$; Scene; $F(3,30)=1.27, p=0.3, \eta^2 = 0.11$). (Also Kruskal-Wallis non-parametric comparisons provided results similar to those of the one way ANOVAs).

4.2.7.4. Case Hypothesis – role of scenario

A statistical examination of the effect of scenario on the number of hypotheses that were generated, could not be conducted, because of the small numbers of hypotheses generated. Participants tended to generate only one hypothesis for each scenario with the exception of Scenario-2, where sometimes more hypotheses were generated.

The most common hypothesis besides murder for scenario-1 was self-harm (suicide), generated by 3 out of 35 participants. In scenario-3, 8 participants suggested assault or not enough evidence as alternatives to murder, while another 16 participants did not suspect a crime had been committed. Scenario-2 led to four main hypotheses: accidental murder, sexual assault and murder, drug overdose and suicide, but these were generated by different participants and none of the participants generated all four hypotheses.

4.3. Discussion

To examine whether CSIs and novices differed in their examination of crime scenes, images of simulated crime scenes were presented and participants were asked to indicate which parts of the scene they thought were relevant to the case. While they were performing the task,

their eye movements were recorded. After the task, they were shown their recorded eye movements and were asked to provide a verbal report about their thought process.

Several expertise effects were found, although they were all relatively small in size. Saccadic amplitudes, dwell times on evidence, state evidence, and fixation durations showed a difference between experts and novices. However, similarities between CSIs and the other groups in how they approached evidence selection were found. In contrast to the relatively small effects of expertise, scenarios strongly influenced participants' eye movements and evidence selection during the tasks. These data suggest that the stimulus layout (which varied across images) more strongly influenced participants' eye movements than their expertise in the crime scene investigation process. In terms of top-down and bottom-up processing, the bottom-up factor tested here (image layout) therefore seems to have a larger effect than the top-down factor tested (expertise). However, this does not mean that bottom-up factors necessarily have larger effects than top-down factors. There may be bottom-up factors that contribute less to the patterns of eye movements than certain top-down factors, and more research is needed to determine the relative contribution of other bottom-up and top-down factors.

The study used photographs of simulated scenes, allowing for high precision eye tracking. Although it may be argued that photographs are unlike actual crime scenes, it must be noted that photographs are highly important in the crime scene investigation process. Photographs are used to document all aspects of the scene. These are taken before the crime scene investigations start, but also during the process to document the investigation of the scene. The photographs are taken following highly standardized procedures and rules (Fisher & Fisher, 2012). Although 3D imaging is becoming more popular in the context of the crime

scene investigation process (DeLaurentis, 2009; PoliceOne, 2015) mainly due to economic constraints, 2D imaging is still the routine standardized procedure. Photographs are therefore ecologically valid materials to study expertise effects in the crime scene investigation process. This is unlike other domains, where the use of photographs is less common, such as in surgery (Law, Atkins, Kirkpatrick, & Lomax, 2004; Wilson et al., 2010; Wilson et al., 2011) or aviation (Fitts, Jones & Milton, 2005), where photographs as stimuli are less valid to examine expertise effects. Although the photographs were taken following procedures aforesaid above in order to represent the best naturalistic impact about the scene, there is a documented risk that photographers place objects of interest in the center of the image (MacKenzie, Westwood, 2013; Tseng, Carmi, Cameron, Munoz, & Itti, 2009), which may have subsequently biased participants' eye movements. Possibly, this bias may have reduced expert influences on eye movements.

One of the strengths of the current study was the range of participant groups used, allowing for an investigation of the trajectory of expertise in the crime scene investigation process. Past studies of expertise effects have often considered only two levels of experience (e.g. majority of studies in surgery, Hermens et al., 2013, but see Kocak, Ober, Berne & Melvin, 2005 who used three groups), and it is therefore often unclear whether the transition from novice to expert is a discrete or gradual transition. The use of a control group that was age matched to the CSIs was another strength. Novice and experts normally differ in age, and any differences in eye movements could be either the effect of age or of expertise. Using an age-matched control group allows for the assessment of age effects, that a comparison between junior and senior operators cannot do.

In the present study, I did not find that experts spent less time on the task, in contrast to earlier findings in crime investigations (Baber & Butler, 2012). A possible reason may be the mouse clicking task. The results indicated that experts selected more evidence than novices. Since it takes time to select an object using a mouse click, experts may have taken relatively longer times to complete the task, even when being more efficient at identifying the evidence.

While the crime scene investigation process can, in part, be thought of as a visual search task (the evidence has to be found), there are clear differences with a classical search task. In the search, new objects need to be linked to objects previously located and hypotheses need to be formed about how the objects fit in the overall picture of the crime scene (Baber, 2010). In this process, objects also need to be ignored (not relevant to the assumed crime). Unlike a classical visual search task in which a specific target is at stake, there is no pre-defined clue (e.g., an object differing from other objects in the scene by a certain feature; Treisman, 1998). In this respect, it is proposed that crime scene examination requires both global and focal attention allocation in order to acquire detailed information in the crime scenes and to interpret the relevance of the objects, materials and places that could yield clues for the further investigation. The present analysis makes use of a specific set of areas of interest, and it may be that difference between experts and novices were not captured by these specific areas of interest. This is a general issue with the areas of interest approach, in which a selection has to be made regarding the areas of interest analysed. Although alternative approaches exist, in which the distribution of fixations are compared between groups for the same image (e.g., using the IMAP software; Caldara & Miellet, 2011) such methods have other issues, such as how to optimally correct for the many statistical comparisons made in the process of comparing participants' eye movements.

The CRQA did not reveal many differences in the scanpaths between the different participant groups. This result may have been limited by the three measures that were used to quantify the overlap in scanpaths. More detailed analyzes of scanpaths are available (e.g., Cristino et al., 2010; Dewhurst et al., 2012; Duchowski et al., 2010; Heminghaus & Duchowski, 2006; Le Meur & Baccino, 2013; Mathot, Cristino, Gilchrist, & Theeuwes, 2012), but these are highly computationally intensive, and would therefore be difficult to carry out. In order to make a group comparison, the results of such analyzes would still need to be condensed in a single number per participant pair and pooled across participants, limiting the information gained from such analyzes. A more important limitation of other scanpath techniques is that they do not pick up similar scanpaths that are simply shifted over time. This is where the greatest advantage of CRQA is: It will detect similar scanpath no matter whether they start at the same time. The one CRQA parameter that showed expertise effects was determinism. Interestingly, determinism differences were predominantly found between first and third year forensic students, not between the novices and the experts. The determinism parameter suggest that the ordering and convergence of fixations differ between the two groups, but it is unclear why this difference was found only in the comparison of first and third year students.

The verbal reports were consistent with the mouse data, and showed that CSIs also made more references to the scene, in agreement with findings by Schraagen and Leijenhorst (2001). Interestingly, the differences between CSIs and third year students on this aspect (selected evidence) were larger than between CSIs and first year students. This suggests that skill acquisition in crime scene investigation is not a linear process. This result agrees with findings in other domains that also suggest that people with an intermediate level of expertise demonstrate worse performance than experts and novices, possibly while they are coping

with the complex build-up of expertise (Gegenfurtner et al, 2011). While students in their third year have learned about the theoretical underpinnings of the crime scene investigation process, they have not been able to apply this knowledge in practice. In contrast to the first year students, they may try and apply their theoretical knowledge to the task, and may therefore display worse performance.

The results can be understood in terms of bottom-up processing on the one hand (identifying places and objects and top-down processing on the other hand applying knowledge about crime scenes; Baber, 2010; Lee & Pagliaro, 2013). It is the top-down processing that is thought to be most strongly influenced by experience (Dror, 2011). The data suggests that top-down influences are strongest for the verbal reports. Other aspects, such as selection of objects by eye gaze or mouse clicks, appears to be largely defined by bottom-up processes. These results agree with those found in medical expertise, with top-down effects mostly on reasoning and decision making processes (e.g. Kundel et al., 2007; Nodine, Kundel, Lauver & Toto, 1996).

The mouse selection data may have been biased by the instruction to participants to identify all objects that may contain evidence. This may have led participants to select as many objects as possible trying to avoid missing important objects. This may have led to a ceiling effect, and a reduction of differences between groups. Part of the crime scene investigation process is to know what evidence to focus on, and therefore a more selective approach towards selecting evidence would more clearly reveal differences between experts and novices. One can, however, argue against this limitation by noting that CSIs were more consistent in their selection of evidence than novices, meaning that CSIs did not select any objects found in a crime scene.

The present study did not evaluate “probative value” (i.e., the extent to which evidence is proof) of the selected evidence. I asked participants to indicate whether objects or part of the scene were relevant, but in future studies, I instead may also ask to indicate how strong participants consider the value of the object for the case is. Probative value has been shown to be of importance for generating hypotheses about a case and for linking the scene to a specific suspect (Lee & Pagliaro, 2013). Future studies should therefore examine how experts and novices differ in their judgments of the value of individual evidence. By exploring whether there is an association between the probative value and eye movement data on different evidence types may provide important insights about the perceptual and cognitive processes of CSIs.

The lack of statistical differences between groups may be related to the relatively small sample sizes used. This is a common problem in research on experts, where there is a limited number of experts available (e.g., there are only a small number of surgeons in each hospital; Hermens et al., 2013, for a discussion of this problem in the context of surgery). The only way to recruit larger numbers of participants when experts are involved, is by conducting studies across multiple sites. This requires collaboration with many police forces, and possible with research teams across universities. As a possible alternative approach, results may be obtained in smaller studies, and later combined in a meta-analysis. Such an approach, however, requires studies to use very similar approaches, so a method of standardising the studies would be required (it is for these reasons that the results in surgery could not be submitted to a meta-analysis, as paradigms used so far were too diverse; Hermens et al., 2013).

While participants were asked not to talk to their colleagues about their participation in the study to others, it is not clear whether participants followed this instruction, and it may be possible that some of the participants were already informed about the simulated cases before taking part in the study. This may have been a factor in the limited group differences.

4.4. Conclusions

In the present study, consistencies in eye movement data between participant groups suggested that bottom-up influences complement top-down factors. Although selection of evidence showed similarities between groups, the differences in references generated for evidence in verbal reports suggested that investigators selected the evidence for its value for future analyzes rather than its visual properties (see also, Baber & Butler, 2012). This study presented a novel approach to examining the importance of perceptual and cognitive processes in examining crime scenes while viewing crime scene images, the findings did not show clear effects of expertise. It is now possible and easier to use the eye tracking methodology in real world settings with the mobile eye tracking devices. A further study using mobile eye trackers (see Chapter 6) aimed to provide in-depth analysis of CSIs' perceptual and cognitive processes in a case work (like) environment, and consequently lead to more generalizable outcomes.

Chapter 5. Comparing expert and non-expert CSIs in a Change Blindness Experiment

5.1. Introduction

Observers make eye movements to shift their fovea (the most sensitive part of the retina of the eye) towards objects of interest for detailed processing. These shifts of gaze have the further benefit of selecting objects for processing, because due to the limited capacity of the cognitive system, not all incoming visual information can be processed in full detail.

These eye movements will be measured in the context of a change blindness paradigm. Change blindness is one of the demonstrations of the limited processing capacity of the brain. It refers to the finding that observers struggle to detect changes in a visual scene, when subsequent presentations of the scene are separated by a transient event, such as a blank interval (Simons & Levin, 1997), although change blindness has been reported in the absence of visual disruption (Simons, Franconeri & Reimer, 2000). The phenomenon is strongly related to intentional blindness, where observers fail to notice often salient events when engaged in another attention demanding task. For example, the appearance of a gorilla in a movie clip was often undetected when observers were engaged in counting how often one of two basketball teams were passing the ball. Likewise, radiologists were found not to detect the image of a small gorilla in a sequence of radiological images, when they were focused on detecting medical abnormalities in the image (Drew, Vö & Wolfe, 2013).

One of the most common paradigms in change blindness studies is the flicker technique (Rensink et al., 1997). In this technique alternating presentations of the image (with and without a change) are separated by a mask, often a blank screen (see Figure 5.1). Participants are asked to detect or localize the change in the image as quickly as possible. Typically, images are each presented for around 240 ms, separated by a blank for around 80 ms (Rensink et al., 1997), although other timings have been used (e.g. Werner & Thies, 2000).

The change blindness technique may be particularly suited to pick up subtle differences between perception and cognition process of experts and novices. Expertise in a domain may be highlighted by a sensitivity to objects that are “out of place” or knowing what a typical scene should look like. Experts may therefore, under the limited presentation times in a change blindness display, be quicker to find parts of the scene that have been manipulated. For example, strong expertise effects were observed in a change blindness paradigm in American football players, specifically for images related to the game (Werner & Thies, 2000). Professional football players were faster and more accurate to detect domain-specific changes (e.g., adding a second ball to the scene) than controls, but they were not better at detecting changes that were not related to the game (e.g., a colour change of the referee's glove, or a change of the location of a player's shadow) or at detecting changes in control scenes (e.g., traffic scenes). Similar expertise effects on change blindness were found in the context of solving physics problems (Morphew et al., 2015). Participants were presented physics diagrams with domain relevant and irrelevant changes. As before, experts were more accurate and faster in detecting domain relevant changes.

Studies indicate that, to successfully detect changes in change blindness paradigm, focused attention should be directed to the relevant location. Alternating changes need to be encoded

and compared, and the change needs to be recognized (Simons, 2000). Familiarity and relevance of the object and expertise in the task can have an influence on detecting the changes, but none of these effects completely eliminates the change blindness effect (Feil & Mestre, 2010).

An alternative approach to study change blindness (the flicker paradigm), called “one-shot paradigm”, involves presenting the original and altered images only once (separated by a blank screen or blink). Participants perform a secondary task on the stimuli and are often not made aware of the changes in the stimuli. An example of this approach is the study on the importance of conceptual knowledge and expertise in physics (Feil & Mestre, 2010). Participants (experts and novices) were presented with two identical diagrams, with a small change made while observers looked away from the diagrams for a moment. Experts more often reported the change, but only when the change affected the answer to the problem.

Both paradigms were used by Beck et al. (2013) to study the effects of training and knowledge on visual working memory in radiologists. As before, experts were found to be more sensitive to domain-related changes, but not to domain-unrelated changes. The underlying reason for this finding could be a modulation of attention by expert information retained in long term memory (Beck et al., 2013). Training change detection has been shown to be possible, but only for the specific domain where training took place (Gaspar, Neider, Simons, McCarley & Kramer, 2013).

In the forensic domain, Allison and Adams (2012) found no expertise effect in change detection in monochrome fingerprint images, but viewing strategies differed between novices and experts. Change detection has also been applied to study eyewitness memory. These

studies used video clips in which a detail was changed between cuts (e.g., a change in the driver of a car; or a change in the burglar in a video of a burglary, Davies & Hine, 2007). Instructions to the participants about whether to pay close attention to the details of the video had an effect, with better change detection with an instruction to pay attention to details. Overall, change detection, however, remained low (at around 39%). These effects were unaffected by how serious the crime was (burglary of \$500 versus burglary of \$5, Nelson et al., 2011). This latter study also found very low change detection recognition rates (5%). Finally, in a study with police officers and novices (Smart, Berry & Rodriguez, 2014), similar levels of change blindness were found in both groups in detecting a change of identity of a driver in a mock traffic stop video. While change blindness was modulated by expertise in other domains than forensics (e.g., football), forensic demonstrators of expertise effects on change blindness have so far limited results.

Such a lack of expertise effect on change blindness is important. Noticing changes in a scene is a crucial skill in crime scene investigation. Especially in serious cases, different crime scene investigators are involved in the crime scene examination process. This creates the problem that the presence of these various investigators increase the risk of disturbing the scene and influencing available evidence, for example by unintentionally moving traces. Although all the steps of investigations are documented and well-recorded, crime scene officers are expected to be aware of such undesired risks of change in the location of traces/objects. With this respect, it can be difficult for CSIs to notice changes in a visual scene as it may be in a crime scene. Failure to detect changes means that interference with a scene may not be detected, which can be a crucial failure in linking a perpetrator to a crime. If expertise effects in crime scene investigation are similar to those involved in football and radiology, expertise is expected to influence change detection. If expertise effects in crime

scene investigation are more similar to those involved in other areas of forensics (fingerprint matching, eyewitness memory), performance is not expected to be modulated by expertise, but scanning patterns may be.

The present experiment will therefore test whether the expertise effects on change blindness found in football, problem solving and radiology extend to crime scene investigation, or whether expertise effects in this domain is more similar to the results found in other forensic domains (fingerprint matching, perpetrator recognition). Just the previously conducted experiment, eye tracking will be used during the task, so that the study can verify earlier results suggesting that eye movement patterns in a change blindness paradigm may differ between experts and novices (Allison & Adams, 2012).

5.2. Method

5.2.1. Participants

The same CSIs and age-matched novice participants of Experiment 1 (Chapter 3) participated, with the exception of one male novice participant, who was no longer available. This participant was replaced by four new novice participants. The study, therefore comprised nine experts ($M=40.8$ years, $SD=9.1$, 5 female) and eleven novices ($M=38.27$ years, $SD=7.7$, 7 female).

5.2.2. Ethics Statement

The Ethics committee of the School of Life Sciences of the University of Lincoln approved the study protocol. Participants all provided consent with written instructions and were verbally debriefed at the end of the study.

5.2.3. Materials

Photographs of simulated crime scenes and scenes all from Lincolnshire were presented in a flicker paradigm on a computer monitor. Crime scenes were simulated in a range of environments. To create the stimuli for the study, for each photograph, an alternate version was created by digitally editing the images (using Photoshop), for example, by changing objects or adding objects to the original scene. Due to the diverse nature of the scenes and difficulty to define relevant and irrelevant changes for the domain (e.g., the removal of a knife in the image could be relevant to the case, but irrelevant in another case), comparisons were made for the different categories of images (crime scenes versus Lincoln scenes) rather than between individual changes. The location of the changes in the images was counterbalanced by dividing the overall stimuli into four regions (top-left, top-right, bottom-left, bottom-right) and ensuring that changes happened equally in each of these quadrants. The two versions of each image were then combined into a movie clip in which they were alternated with an intermittently presented blank screen to avoid apparent motion cues (see Appendix B for change information on each image). Each image in the sequence was presented for 500 ms, separated by a blank (white) screen (all white) for 100 ms. The entire sequence of alternating versions of the image separated by blank screens lasted for 30

seconds. This limited duration was used to restrict the overall time participants were required in the study.

The total set of stimuli consisted of simulated crime scene colour images and 15 real scene colour images from all around Lincolnshire (downloaded from the internet / taken with a camera). These latter scenes included scenes of Lincoln, with houses, streets, traffic, historical buildings and cars. All images were rescaled to fit a monitor with a maximum screen resolution of 800 x 600 pixels resolution with each picture image subtending 640 x 480 pixels with a white surrounding background border. The image area was equal to approximately 17 by 13 degrees of visual angle against a white background.

5.2.4. Apparatus

An Eyelink 1000 desk mounted system was used to record eye movements at a sampling rate of 1000 Hz. A 22 inch LCD computer monitor was used, set at 1680x1050 pixels spatial resolution and 59 Hz refreshment rate. Participants were seated at a viewing distance of 60 cm from the computer monitor with their head in a head and chin rest. The experiment was set up using Experiment Builder. Eye movement parameters were extracted from the raw eye movement data (horizontal and vertical eye gaze position) using the Data Viewer software. Participants provided their responses using a MS Sidewinder gamepad.

5.2.5. Procedure and Design

At the start of the experiment, the eye tracker was calibrated using the default nine point calibration procedure. This involved the participant fixating a series of fixation targets

presented on a three by three grid on the computer screen. Calibration was considered successful if the nine randomly illuminating calibration points were aligned with the same three by three grid on which they were presented. Such a calibration corresponds to a reported accuracy of 0.5° in visual angle. If the recorded points were not aligned with the three by three grid, calibration was repeated.

The experiment was then started. On each trial participants were presented with one of the video clips of alternating versions of a scene. Participants were instructed to press one of the buttons on the gamepad as quickly as they detected the change. The display terminated when the participant pressed the button. They were then asked to report the observed change verbally to the experimenter, who then recorded whether the change was detected correctly.

Participants completed an initial practice session with feedback from the examiner to ensure they understood, and to get familiar with the eye tracker. This practice session also provided the opportunity to ask questions. The practice session contained two videos from each type of stimuli (crime scene and Lincolnshire scene), which were not used in the main experiment. Piloting was also carried out by the experimenter prior to commencing the experiment to ensure that digital manipulations were not immediately obvious or problematic.

A total of 14 experimental videos were presented to the subjects in a fixed order (examples in Figure 5.2 and Figure 5.3). In between trials, a drift correction was performed (involving the participant fixating a central fixation target followed by a key-press to confirm fixation) to avoid drift in the eye movement recordings throughout the experiment. Crime scene investigation trials were completed first, followed by the Lincolnshire trials. Participants' verbal answers were written down by the researcher, without providing feedback on the

accuracy of the response. Response times were automatically extracted by the stimulus presentation software created using Experiment Builder.

After completing the trials, participants completed a demographic questionnaire. They were then debriefed, and thanked for their participation. In the debriefing process, participants were asked unofficially about whether they had any knowledge on change blindness phenomenon before the experiment. Out of 20 subjects who participated, only 2 claimed to know about the change blindness phenomenon before the experiment. However, both participants were unable to give details about it. To facilitate participant recruitment, the testing of the CSIs took place the unit's headquarters, whereas the novice participants took part in the Eye Tracking Laboratory at the University of Lincoln.

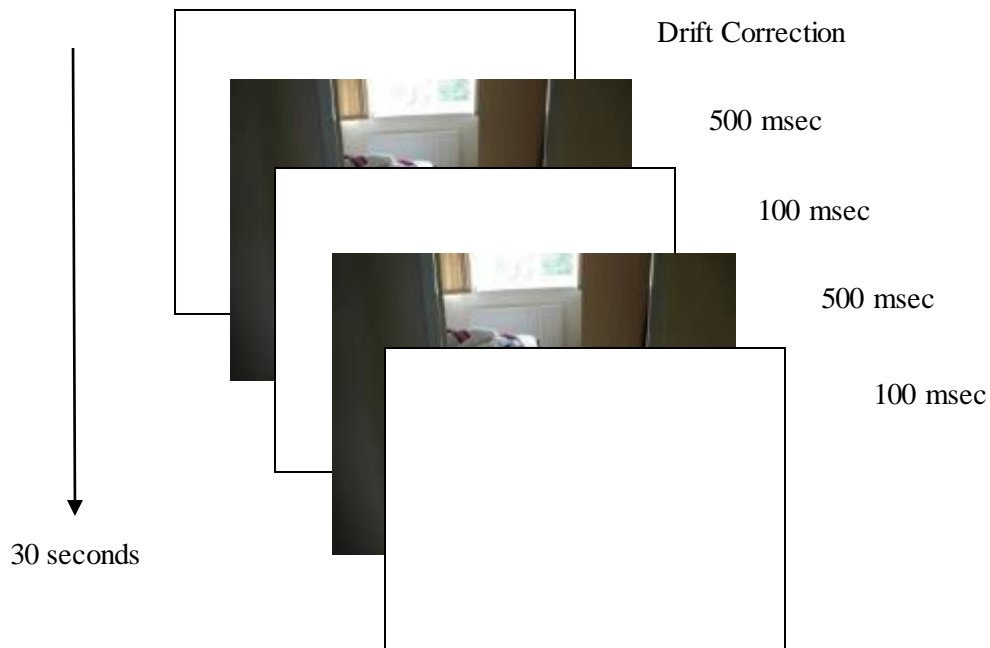


Figure 5.1. The stimulus sequence. Trials started with a drift correction, in which participants fixated a centrally presented target during which the experimenter pressed a key to confirm fixation. Drift correction was followed by the sequence of alternating versions of each scene, separated by blank screens until participants indicated they detected the change, or for 30 sec.

5.2.6. Data analysis

Two measures were extracted: The accuracy of the response and the time needed to find the change. For statistical comparisons of these measures between CSIs and novices, independent samples t-tests were used. Videos were presented for 30 seconds only. Therefore, if participants did not detect the change, the answer was recorded as a miss. Only response times and eye movement characteristics for correct trials were analyzed. Performance on one Lincolnshire scene was poor for almost all participants, and data from this scene were therefore excluded from the analysis. No additional filtering was applied to the eye movement data.



Figure 5.2. Example of a simulated crime scene change. The original image is on the left.



Figure 5.3. Example of a neutral scene change (railway crossing in Lincoln). The original image is on the left.

5.2.7. Results

5.2.7.1. Performance Data

Accuracy

Figure 5.4a plots the accuracy for CSIs and novices for the crime scenes, Lincolnshire scenes, and the overall performance. Performance was generally high (around 75%), although there was quite a bit of variation across participants (accuracy ranging from 35% to 100%). Once a detection was reported, it was mostly correct (CSIs: 5% incorrect for crime scenes, 7% incorrect for Lincolnshire scenes; novices: 3% incorrect for crime scenes, 1% of incorrect for Lincolnshire scenes). Importantly, the results did not show an effect of experience on overall change detection performance (CSIs: 76%; novices: 77%; $t(18)=-0.39$, $p=0.7$, $d=-0.18$).

Past studies, however, have suggested that superior change detection performance in experts is only found in domain relevant trials. A subsequent analysis therefore split the data according to stimulus type. In contrast to previous findings, no effect of expertise was found on domain specific trials either (simulated crime scenes, CSIs: 77%; novices: 76%; $t(18)=0.11$, $p=0.91$, $d=0.05$). In agreement with earlier studies, no expertise effects were obtained for domain unrelated scenes (Lincolnshire scenes, CSIs: 76%; novices: 79%; $t(18)=-0.54$, $p=0.6$, $d=-0.24$).

Response Times

Figure 5.4b plots the response time on accurate trials (change detected within 30 seconds and change accurately indicated). As for the accuracy, no expertise effects were found on response times for the simulated crime scenes (CSIs: 9.07 sec; novices: 9.81 sec; $t(18)=-0.82$,

$p=0.42$, $d=-0.37$), or for the Lincolnshire scenes (CSIs: 8.51 sec; Novices: 8.54 sec; $t(18)=-0.37$, $p=0.97$, $d=-0.01$). No expertise effects were found across scene types either: (CSIs: 8.92 sec; novices: 9.19 sec; $t(18)=-0.34$, $p=0.74$, $d=-0.15$).

5.2.7.2. Eye Movement Characteristics

As one past study in forensics (Allison & Adams, 2012) suggested that scanning patterns may differ between experts and novices in the absence of any performance differences, the eye movement data was analyzed to explore the attentional strategies of the two groups of participants. First, global eye movement characteristics were compared (number of fixations, fixation durations), followed by second, a comparison of dwell times on the changed object. Figure 5.4(c-e) plots the eye movement data for CSIs and novices for the crime scenes, Lincolnshire scenes, and the overall performance

Fixation Durations

Interestingly, a trend towards longer fixation durations in CSIs was observed (Figure 5.4c). Marginally significant differences in fixation durations were found for simulated crime scenes (CSIs: 392.9 ms; novices: 356.5 ms; $t(18)=1.81$, $p=0.086$, $d=0.82$), for neutral scenes (CSIs: 379.1 ms; novices: 349.5 ms; $t(18)=1.79$, $p=0.09$, $d=0.81$) and across both types of scenes (CSIs: 386.5 ms; novices: 353.2 ms; $t(18)=1.89$, $p=0.075$, $d=0.85$).

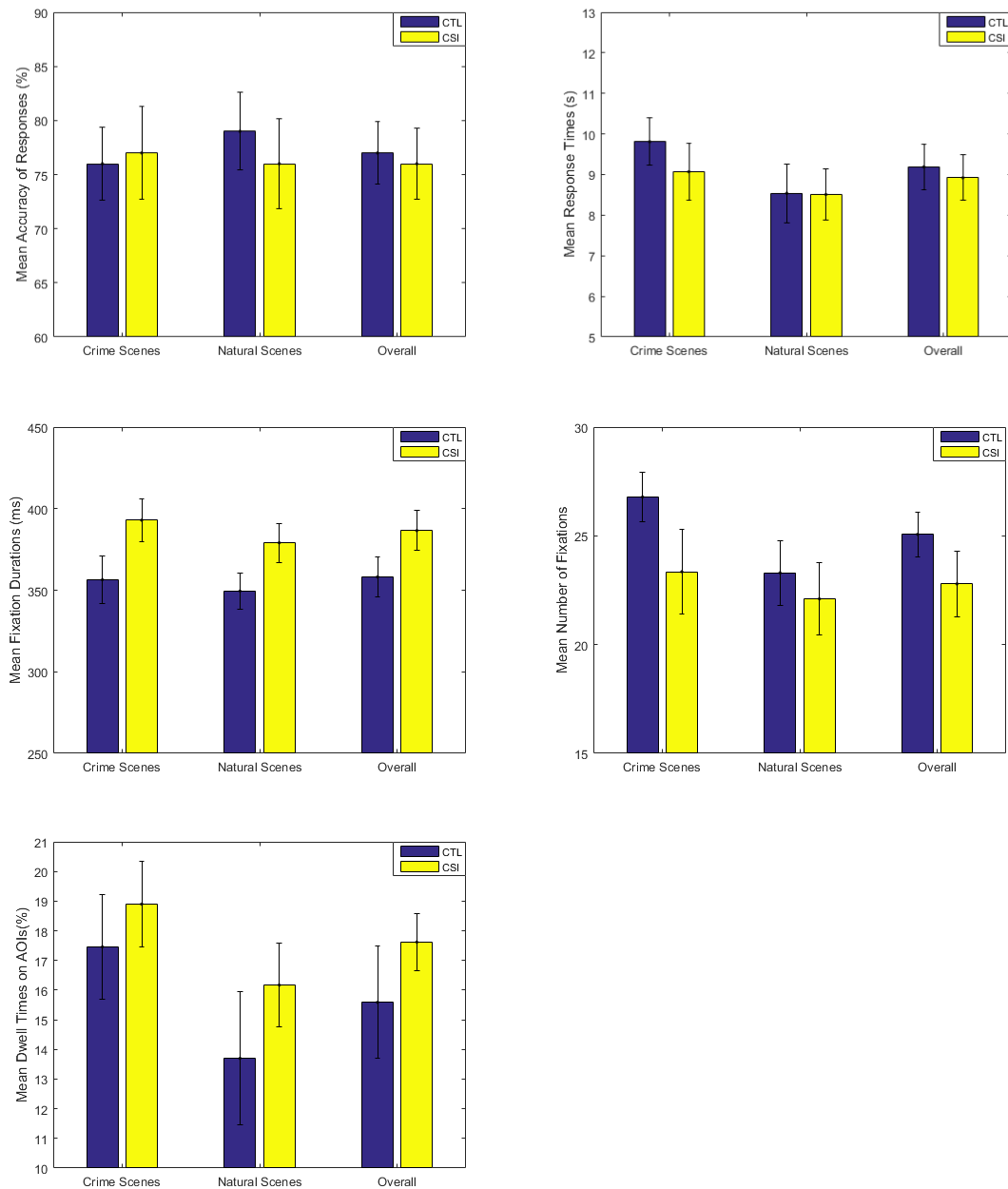


Figure 5.4. Results of Experiment. a) Accuracy of responses (change detected within 30 seconds and correct change indicated). b) Response times on accurate change detection trials. c) Fixation durations on accurate change detection trials. d) Numbers of fixations on accurate change detection trials. e) Dwell times towards the changed object or material on accurate change detection trials for experts and novices (in different coloured bars, see legend, CTL=Control, CSI=Crime Scene Investigators, N=20) across the two stimulus types (left two groups of bars) and across all scenes (right two bars). The error bars show the standard error of the mean across participants.

Number of Fixations

There was a slight trend towards CSIs making fewer fixations during the trials (Figure 5.4d), but none of the comparisons with novices on this eye movement measurement revealed a significant effect (for crime scenes; CSI: 23.4, novices: 26.8; $t(18)=-1.61$, $p=0.13$, $d=-0.71$; for neutral scenes; CSIs: 22.1; novices: 23.3 ; $t(18)=-0.54$, $p=0.6$, $d=-0.24$; across scenes; CSIs: 22.8; novices: 25.1; $t(18)=-1.29$, $p=0.21$, $d=-0.57$).

Dwell Time Percentages

Dwell times on changed objects or materials tended to be longer for CSIs (Figure 5.4e), but variations between participants were fairly large and none of the comparisons with novices yielded a significant effect of expertise (crime scenes; CSIs: 18.9; novices: 17.5; $t(18)=0.6$, $p=0.56$, $d=0.27$); neutral scenes; CSIs: 16.2; novices: 13.7; $t(18)=0.88$, $p=0.39$, $d=0.41$; across all scenes; CSIs: 17.6; novices: 15.6; $t(18)=0.91$, $p=0.38$, $d=0.42$).

5.3. Discussion

Observers are consistently poor at detecting changes in scenes when they are separated by a blank or a moment of inattention (Simons & Levin, 1997). Several studies have suggested that expertise may aid in overcoming this poor change detection performance, specifically for domain specific changes (Beck et al., 2013; Feil & Mestre, 2010; Werner & Thies, 2000), although other studies suggested no effects of expertise (Smart et al., 2014) or only on scanning patterns during change detection (Allison & Adams, 2012). The present study examined how expertise in crime scene investigation influences change detection in simulated crime scenes and domain unrelated scenes (Lincolnshire scenes) and whether scanning pattern during change detection are influenced by expertise. The results suggest no

effect of crime scene expertise on change detection performance, neither for domain related and for domain unrelated scenes. Experts showed a marginally significant trend towards longer fixation durations while inspecting the scenes, and the other eye movement measures considered (numbers of fixations, dwell time on the changed object or material) had non-significant expertise trends in the same direction. The findings therefore suggest that CSI expertise does not reduce blindness to changes, even for domain specific changes, but does seem to modulate scanning patterns involved in the task.

Performance was relatively high. In the 30 seconds interval that participants were given, around 75% of the changes were detected. This high performance was unexpected as changes made to the stimuli were rather subtle (e.g., the removal of the knife under the chair) and the stimulus timing was similar to that used in past studies. Performance, however, was still far from ceiling (100%), and the absence of an expertise effect can therefore not be explained from a ceiling effect. Variability across participants was, however, high. Given the general problem of recruiting large numbers of expert CSIs, this variability may have been a contributing factor in why no expertise effects were obtained. Trends for accuracy and response times for domain related (crime scene) changes were in the predicted direction, and it may therefore be the case that increasing the sample size would lead to significant effects of expertise. This may be achieved, in part, by increasing the size of the control group, and should be considered in future studies.

The present results are in contrast with previous expertise effects on change blindness in football (Werner & Thies, 2000) and radiology (Beck et al., 2013). A possible reason could be that expertise in sports and radiology rely more strongly on perceptual skills than crime scene investigation. Scenes and images in radiology and football are also more restricted in

what objects can appear in the scene. The task is also more specific in football and radiology, compared to the crime scene investigation process. A possible counter argument, however, may be that performance in the present study was not considerably different than the previous studies, for example in Werner and Theis (2000) who found a 78.9% accuracy across 40 seconds search intervals. If the lack of expertise effects would have been due to larger variations in the visual aspects of the images used, one would also expect lower performance. The level of performance, however, is linked to the saliency of the change, and the amount of clutter in the scene, and overall performance may therefore not be linked to expertise effects.

A stronger trend towards expertise effects was found in the eye movement patterns. CSIs had longer fixation durations, had fewer fixations and tended to fixate the change for longer than the novices. These findings are in line with the work by Allison and Adams (2012), who found no change detection performance differences between experts and novices, but did observe differences in scanning patterns between the two groups during the task. Together with their results, the present results suggest that eye movements during the task may be unrelated to the ability to notice changes in a visual scene, and instead, serve a different purpose. The difference in eye movements might reflect a process of encoding the acquired information or to take into account different aspects of the scene (e.g., regarding the relevance of the objects in the scene). In this respect, the results differ from those found in detecting misdirection in magic, where eye movements between participants who understood how the trick was performed did not differ from those who were misled (Kuhn & Findlay, 2010; Kuhn, Tatler & Cole, 2009; Kuhn, Tatler, Findlay & Cole, 2008). The present results are in agreement with the findings by Smart et al. (2014), who did not observe expertise effects when comparing police officers and lay people.

The study is not without its limitations. Only 14 pairs of images were used, although it should be noted that this number is not much different from the numbers in previous studies that did show expertise effects (10 images for each category in Werner and Theis, 2000, and 14 in Beck et al., 2013). The design of the studies was also relatively similar, except for the use of monochrome images in Beck et al (2013), which may explain why their study showed relatively high performance overall. Also, practice with the task seems to improve performance, and such practice effects might be stronger if the images are more similar (e.g., radiography images or football pitch image). The images in the present study were quite diverse, which may have reduced practice effects, although it is not immediately clear how this would affect expertise effects. In the present study, we treated each crime scene change equally. However, changes in crime scenes vary in their importance to the crime scene investigation process. Future studies should therefore examine whether expertise effects influence change detection of objects of different importance to the scene differently. The direction of such effects may not be immediately obvious. On the one hand, experts may be better at detecting important scenes. On the other hand, since important changes may be obvious, experts may instead differ from novices in how well they detect less important changes.

In the present study, a possible reason for the absence of significant differences may have been the relatively small sample size. Previous studies employed considerably larger sample sizes (48 experts and novices in Werner et al., 2000, 57 experts, intermediate and novices in Beck et al., 2013) and 29 participants in Galpin et al., 2009). The sample size in the present study, because only one of the police forces approached, agreed to take part. If more police forces would have agreed, a large sample would have been possible, as shown by Allison and Adams (2012) who combined 9 experts from one department and 13 from an another

department. To achieve a power of 50% with a large effect size ($d=0.8$), both groups should include 14 participants (calculated using G-Power using priori analysis, Faul, Erdfelder, Buchner, & Lang, 2009). This suggests that lack of statistical power may have been an issue for the present study.

Eye movements were tracked in Experiment 2, allowing for a comparison with the findings by Allison and Adams (2012), who found differences between experts and novices. Eye movements showed a tendency in the same direction as in Allison and Adam (2012), with more experienced CSIs showing a broader scanning pattern of the scene. While change blindness accuracy does not seem to provide a sensitive measure of skill, eye movements during change blindness may play a role in skill assessment. Similar dwell times on the change locations (AOIs) were found across groups. This can be understood from the similar performance between groups: Once the changing object was found, people tended to focus on this area.

5.4. Conclusion

In contrast to previous studies, the present study failed to find a reduction of change blindness from expertise with crime scenes. Eye movements measured during the task suggest different scanning strategies for expert CSIs, but these effects need to be confirmed in a larger sample. As the present study only looked at change detection in images, future studies should consider change detection in actual or simulated crime scenes.

Overall, Experiments 1 and 2 found limited effects of expertise on performance and eye movements in crime scene investigation. A possible reason could be the use of static images

as stimuli, which are quite unlike actual crime scene investigation. Studies in surgeons have suggested that expertise effects on eye movements can only be found when actively engaged in (simulated) surgery, and much less so when watching surgery (for a discussion, see Hermens et al., 2013). Chapter 6 will therefore employ an actual (simulated) crime scene investigation and will compare volume (less experienced) and senior (more experienced) CSIs using mobile eye tracking.

Chapter 6. Eye Movements in Volume and Senior Crime Scene Investigators during Active Exploration of a Simulated Crime Scene

6.1. Introduction

Perceptual and cognitive processes of interest include observational skills, induction abilities and technical understanding (Williams, 2004, as cited in Ludwig et al., 2012). Research on these factors, however, is limited. This lack of research may explain why CSIs are often regarded as the mere collectors of evidence in a crime scene (Ludwig et al. 2012). Likewise, CSIs (and other forensic staff) are hardly aware of the contributions they make in an investigation. (Ludwig et al., 2012). That is why, CSIs were described as the “backroom boys” in overall investigation process (Wilson-Kovacs, 2013). As a consequence, the value of the work conducted by CSIs has not always been appropriately appreciated (Cole, 2013).

In order to conduct an effective investigation, domain-specific knowledge, experience, skills and attitudes play a major role in addition to the technologies that aid the investigation (Kelty, et al., 2011). A full grasp of cognitive process including perception, reasoning, decision making and training, experience and coordinated team spirit is essential for quality crime scene investigation (Lee & Pagliaro, 2013). In the crime scene examination process, cognitive processes play an important role, in addition to the motor skills (collecting and packaging traces; Lee & Pagliaro, 2013). The crime scene examination process involves different stages, as explained in previous chapters. Because it is impossible to investigate all stages in the context of this thesis, the focus will be on the initial stages of the crime scene investigation process. This initial stage requires skilled perceptual and cognitive processes to

identify relevant information, utilizing domain specific knowledge, with the goal of interpreting the evidence (Baber, 2010).

By comparing novices and experts, I aimed to examine the influence of expertise on the perceptual and cognitive processes involved in the crime scene investigation process. Chi (2006) suggests two approaches to study expertise effects. The first approach studies performance of well-known exceptional experts in the domain. The second approach compares experts (not necessarily exceptional) to novices. A third method is possible, but this requires more resources. In this third method novices are followed over time while they develop their skills and become experts (Chi, 2006). Because exceptional experts are not readily available and because the time-span of a PhD project does not suffice to follow a novice crime scene investigator until they become an expert, I relied on experienced, but not exceptional CSIs. Because a location for the study near the CSI headquarters needed to be chosen (to increase the chances of having CSIs take part in the study), no novices were available for the study. I therefore used a combination of the first two approaches: I only examined experienced CSIs, but they were not the top of their field.

As in the previous experiment in Chapter 4, eye tracking was used as the main method of studying perceptual and cognitive processes in crime scene investigation while using static images of simulated crime scenes. However, the present study will examine eye gaze of CSIs while inspecting a real-world simulated crime scene. The advantage of using static images, is that the same stimuli are presented to all participants, and that's why eye tracking is highly accurate. The disadvantage is that viewing crime scenes on a computer screen while sitting in a chin rest is unlike the actual crime scene investigation process. Recent studies have suggested that results from screen based eye tracking studies do not automatically transfer to

results in real world eye tracking (Kingstone, Smilek & Eastwood, 2008). Such differences in results were particularly prominent in social attention research, where it was found that participants fixate people when presented on a computer screen. In the real world, however, fixating other people is avoided, unless one aims to engage in a conversation with the other person (Laidlaw & Kingstone, 2015). Differences in gaze patterns were also obtained when tracking goal keepers' eye movements when watching a penalty taker on a computer screen, compared to watching an actual penalty taker (Dicks et al., 2010). As mobile eye tracking is a technique still very much under development, these two differences observed between mobile and screen based eye tracking suggest in the limited number of comparison studies undertaken thus far suggest that mobile eye tracking results may differ from screen based results in many more domains. The present study therefore uses state-of-the-art mobile eye tracking technology to study eye movements during the crime scene investigation process, to better understand how perception and cognition are involved in the process. While mobile eye trackers sample at a lower sampling rate and have poorer spatial resolution of desk-mount eye trackers, the use of mobile eye tracking technology greatly enhances the ecological validity of the results. Data analysis will involve more manual coding than for static eye tracking. Due to the head movements of the participants, the images will change during recording and will differ across participants, and therefore eye movements will be needed to be coded for each video frame and each participant. Summarizing the data is also more complicated, as can be seen from the extensive reports of eye gaze behaviour in earlier applications on the mobile eye tracking technique, studying tea making (Land et al., 1999), sandwich making (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003), and surgery (Wilson et al., 2010).

As in the studies with the static images and desk-mounted eye tracker, this study with the mobile eye tracker complemented the eye tracking with cued retrospective verbal reports. Although only CSIs working in one of the police forces were recruited, expertise effects could still be studied in line with the level of CSIs in the Police Force. Approximately half of the CSIs were volume CSI (lower level of expertise), while the rest were senior CSIs (higher level of expertise). The year of expertise for each of the CSIs was used to study expertise effects.

Besides studying the effects of expertise, the contribution of advance information to the crime scene was investigated by creating two different scenarios for the participants. As stated in Chapter 1, the prior information has influence on the decision making process in crime scene examination (Helsloot & Groenendaal, 2011). Also a recent study of Van den Eeden, de Poot and van Koppen (2016) revealed that CSIs mostly based their decisions on background information even though the same traces that were important to the case were found. This contextual information has a potential for bias on CSIs decisions on the overall case assessment. With the help of eye tracking device, in addition to the assessment in the crime scene, the influence of prior information on perception process of the CSIs was explored in the study using two different scenarios. These different scenarios only differed in the cover story provided to the participants before entering the same crime scene.

6.2. Method

6.2.1. Participants

Fifteen CSIs (CSIs; 8 Senior CSIs, 7 Volume CSIs; 4 female) were recruited with varying levels of expertise (years of experience) from a regional UK Police Force. Participation in the study was approved by the Area Forensic Manager. The mean age of senior CSIs was 43.7 years. These CSIs had on average 15.4 years of experience. Volume CSIs were on average 41.3 years and had 7.3 years of experience. The difference in experience between the two groups was statistically significant, ($t(13) = -2.34, p = 0.036, d = 0.3$).

Recruitment of the participants was achieved as explained for the previous studies, by getting in contact with the Area Forensic Manager of the Police Forces. After permission for the study was obtained, an internal e-mail for volunteers was sent out by the manager to possible candidates for the study. Nominees who were interested in participating in the study were scheduled for the experiment by the Manager on specified dates and times in July 2015.

6.2.2. Ethics Statement

The Ethics committee of the School of Life Sciences of the University of Lincoln approved the protocol. Participants all provided written consent, and were provided with written instructions, and were verbally debriefed at the end of the study.

6.2.3. Materials

Inside the kitchen area of a local church near the police headquarters, an assault scenario was simulated by placing items on floors and furniture in the room (see Figure 6.1). The scenario was designed with the help of an experienced CSI.

Two short background stories about the case were created to examine the influence of advance information on search and inspection strategies (see Table 6.1). The first story told about a missing person, whereas the second story informed the investigator that a person was found, who has been taken to hospital. The scene for the two stories was identical. Stories were randomly assigned to the participants.



Figure 6.1. Photographs illustrating the simulated crime scene set in the kitchen of a local church close to the police headquarters.

Table 6.1.

The two scenarios which were presented to the participants before entering the scene

Scenario-1

Nicolette Kiddom's mother from Canada called the police (July 16th, 10 22 am) and reported that she couldn't have reached her daughter for more than a day. Her mother didn't know anyone from the city and was afraid for her daughter's life. The girl was living alone in her flat. When they last talked two days ago, she mentioned about a guy, named Carr, who was a student at the same class.

Officers went to her address. The door was locked. With the help of the landlord, they opened the door. Nobody was at the house. The kitchen part of the flat was in a mess and there were some stains on the floor. The place cordoned off, the detectives on duty informed. Crime Scene requested.

Scenario-2

Morning shift supervisor of the Shop X called the police (July 16th, 10 22 am) and informed that an employee, Nicollette Kiddom, has not come to work in the morning. What they knew, she was living alone in her flat and she had no close relatives in the city that they could contact and ask for her. Her phone couldn't have been reached. He knew that yesterday, when she was at work she seemed very upset.

Officers went to her address but the door was locked. With the help of landlord, the door opened and they found her lying on the floor in front of the TV hardly breathing. She was immediately taken to hospital. The kitchen part of the flat was in a mess and there was a bad smell. The place cordoned off, the detectives on duty informed. Crime Scene requested.

6.2.4. Apparatus

Eye movement data were recorded using a head-mounted Positive Science Ultraflex Headgear eye tracking system which had a monocular camera to record the right eye's movements while utilizing the combined pupil and corneal reflection signals with a sampling rate of 30 Hz. Eye movement data were combined offline with recordings from a scene camera taking images at the same rate from the observer's point of view. The eye tracker was controlled by a small laptop attached to a backpack, which also stored the collected recordings. The eye tracker was mounted on a head frame worn by the participant, which also included infrared emitter that illuminated the participants' eye to induce the corneal reflection from the pupil. The eye tracker allowed free head movements and navigation around the scene.

6.2.5. Procedure and Design

Before the experiment, participants were introduced to the eye tracking equipment and were explained the calibration procedure. The experiment started with task instructions and an initial practice session, during which the researcher provided assistance to ensure that the participants got familiar with the eye tracker and understood the task requirements. This practice session also provided the opportunity to ask questions and arrange the position of the eye tracker and backpack for the study. The practice session involved a walk in the corridor while wearing the eye tracker and the backpack. It also included an explanation about the cued retrospective think aloud procedure. This procedure, which would take place after inspecting the scene, required participants to verbally express their thought processing while they watched the movie clip from the participants' point of view taken during the

inspection of the scene. After finishing the practice session, participants were given the case scenario.

After reading the case scenario, the eye tracker was calibrated using a five-point calibration process (involving participants looking at the corners and centre of a checkerboard, mounted vertically, as instructed by the examiner). Participants then entered the simulated scene. After finishing the task, the calibration process was repeated to ascertain that the calibration was not altered during the task.

The inspection part of the study was started once participants entered the simulated scene through the door and participants started their search and assessment process. No time limit was given for the scene examination. The experimenter waited outside the room and was available for assistance if required.

After completion of the task and removal of the eye tracker, participants completed a short demographic questionnaire. Participants then performed the “retrospective think-aloud” protocol during which they watched their own eye movement recordings of the task on the laptop of the eye tracker. A recorder (smart phone) was used to record the report, which was analyzed offline. In the protocol, participants focused on three questions:

“Question 1. While you are watching your eye movement recording on the computer screen, provide details about what you were thinking. Explain your thought process.

Question 2. What is your opinion/s about what happened in this scenario?

Question 3. How do you prioritise and rank the objects that could be evidence or the places in the scene? What will be the most important things for the scenario?”

When the task was complete, the participants were thanked for their participation. The experiment took around thirty to forty-five minutes to complete for each participant, depending on how long the participant took for the inspection of the scene.

6.2.6. Data Analysis

Analysis of the eye movement data

The eye tracking data were analyzed using the automatic pupil-tracking software “Yarbus” and the coding software “Gazetag”, both of which were provided with the mobile eye tracker. The pupil tracking software Yarbus uses the raw eye tracking signal as input and maps the position of the eye to a reference system which estimates the X and Y coordinates and start and end of fixations. The coding software Gazetag takes the data and video files from the Yarbus software and utilizes the fixation detection algorithm to calculate the duration and frequency of the fixation. The software also assists the labelling (coding) of the fixations using a library created by the experimenter (containing codes, such “glass”, “table”, and “wall” to indicate the objects that participants were looking at). Participants’ eye movement recordings were coded using AOIs. For a fixation to be assigned to an AOI and to be included into the analysis, I required it to last at least 50 ms. This duration was based on the minimum duration suggested and used in the literature (Nystrom & Holmqvist, 2010; Rayner, 1998; Tatler & Vincent, 2008).

AOIs included objects (phone, shoe, flower pot, etc.) and places (door, wall, ceiling, window, etc.) in the scene, such that every area of the scene was covered. Inevitably, AOIs differed in shape and size. In a subsequent analysis, I categorized each of the object AOIs as “Placed” (deliberately set down for the case scenario) or “not Placed” (all other objects). The rationale

for this definition was that the previous experiments suggested that experienced CSIs look more at wider cues in the scene (likely to be “not Placed” objects), and that novices tend to focus more strongly on objects/materials which could be potential evidence (more likely to be “Placed” objects). The classification of AOIs into “Placed” or “not Placed” was made with the assistance of expert CSIs and allowed for a distinction of fixations on objects in where they could normally expect to be (“not Placed” objects) and objects that were “out of place” (“Placed” objects). For example, there was a broken cup in the kitchen cupboard and some fragments of it were on the floor. This “Placed” item (not in its normal place) could serve as an indicator for an assault. Likewise, the upside down slippers or broken phone in the middle of the scene were also “Placed” objects (not in their ordinary places such as the broken cup). A second categorization of the AOIs involved possible points to go inside or to leave the scene (doors, windows etc.), which I named “Entry and Exit Points”, although in the case scenarios the door was mentioned as locked (but it may have been assumed that the people involved had a key to open it). This categorization allowed for a direct comparison with the results by Baber and Butler (2012), who found differences between viewing strategies across different experience levels. Their findings revealed that experienced CSIs tended to spend more time on viewing and examining the entry or exit points. Finally, on the basis of the verbal reports, AOIs were classified as “Possible Evidence” or other objects (there is some overlap with ‘placed’ AOIs). An overview of all AOIs is provided in Table 6.2 and Figure 6.2.

A range of statistics was analyzed and compared them across the two stories and the two groups of investigators (volume and senior). The first measure considered was the total time spent inspecting the scene. With more experience, senior investigators may take less time than their less experienced colleagues. Predictions for the time spent across the two scenarios

are more difficult to make, and I have no a priori reason to believe one scenario or the other will lead to longer inspection times. Second, I examined cumulative dwell times to the different AOIs as a percentage of the total viewing time. Based on earlier results, senior investigators are expected to spend more time looking at the room (walls, floors, ceilings, doors) and less at the objects in the room than volume investigators. Senior investigators are also expected to look more at exits and entries than their more junior colleagues.

While overall time and dwell time provide information about how long people spent in their investigations, these two measures do not indicate in which order they consider the evidence. The two measures were therefore complemented by a CRQA of the eye movement data. This analysis allows for comparisons of viewing patterns across participants, and for establishing whether similar viewing patterns are used, even when separated by a delay. The analysis generates a range of measures indicating similarities between viewing patterns. These are: recurrence, determinism, laminarity. More detail about each of these measures can be found in Chapter-3. As units for the CRQA, I used the above AOIs, which were each given a number to serve as grid numbers in the Fixed-Grid Method for the analysis (Anderson et al. 2013).



Figure 6.2. The layout of the Scene and Coded AOIs, as listed in Table 6.2.

Table 6.2.

Coded AOIs in the Scene.

Number	Name of Coded AOI	Category		
		1 (Placed)	2 (Entry&Exit)	3 (Evidence)
1	Main Door		X	
2	Main Door_Frame		X	
3	Shoe	X		
4	Coat	X		
5	Key	X		
6	Umbrella	X		
7	Book Shelf			
8	Bag	X		
9	Floor			
10	Wall			
11	Ceiling			
12	Windows		X	
13	Sofa -2			
14	Piano & Seat			
15	Heater on the floor			
16	Sofa -1 & Cover	X		
17	Heater on the wall			
18	Slipper_left	X		
19	Phone	X		X
20	Footmark & Newspaper	X		X
21	TV Shelf			
22	Flower pot	X		
23	TV			
24	Locked Door		X	
25	Heater on the wall			
26	Chair -1	X		
27	Note	X		
28	Table & Items	X		X
29	Chair -2			
30	Slipper_right	X		
31	Broken cup_on the floor	X		
32	Blood	X		X
33	Broken cup_on the kitchen	X		
34	Boxes			
35	Kettles & Plug			
36	Glass	X		X
37	Bleach	X		X
38	Hand Wash			
39	Knife	X		X
40	Sink			
41	Bin	X		X
42	Camera			
43	Cupboards			
44	Tap			

The analysis was conducted using custom-built scripts (Perl and MATLAB) and CRQA Matlab toolbox by Richardson, Lopresti-Goodman, Mancini, Kay and Schmidt (2008) and Richardson, Schmidt and Kay (2007). Statistical comparisons involved independent samples t-tests with 'Experience' or 'Scenario' as between subjects factors.

Analysis of the data from Verbal Reports

As indicated above, the final phase of the study involved a verbal report while watching a playback of participants' own eye movements. These verbal reports were manually transcribed and analyzed. Reports were divided into segments to aid the coding process. These segments and the following coding procedure were inspired by the process proposed by Trickett and Trafton (2007). Verbal data was then coded using the same coding scheme as in Experiment-1, originally developed by Baber and Butler (2012). The following coding scheme was used:

- a. Reference to Objects in the scene. Knife, glass, cups etc.
- b. Reference to Scene and Scene's features. General parts of the room, door, wall or type of flooring etc.
- c. Reference to Modus Operandi. Perpetrator's or other activities that inferred facts about the case
- d. Reference to analysis or evidence. Possible explanations for plans for further examinations (e.g. DNA, fingerprints)

In order to ensure the reliability of the coding, a second coder coded 25% of the verbal data segments independently. The first coder was the experimenter, while the second coder was an experienced CSI with extensive training and knowledge about qualitative analysis. As a

measure of inter-rater reliability, Mezzich's K statistic was used and checked for the sufficient agreement, following the example by Eccleston et al. (2001). Disagreements were resolved through discussion.

6.2.7. Data quality

Data from four of the participants (2 Volume CSI and 2 Senior CSI) had to be excluded from the analysis because of the poor quality of the eye movement recordings (e.g., degradation of the calibration due to movement of the device while performing the task, as demonstrated by the calibration check at the end of the recording) of the eye movement data. For the remaining participants in the analysis, on average, eye movements of 72% of the video clip durations could be coded into AOIs. Data loss for the remaining sections occurred due to poor light conditions for tracking, limits of the eye tracker's field of view or blinks. I here relied on common sense, as there does not seem to be an established protocol for how to deal with missing data in the mobile eye tracking literature. For example, when the eye tracker indicated that a CSI looked just below the knife while holding it in their hands, I coded the eye movement as belonging to the knife, as it can normally be assumed that an eye movement recording near the position of the knife while holding the knife, was actually targeted towards the knife.

6.2.8. Results

The results reported here are based on data from 6 CSIs (3 Senior, 3 Volume) in Scenario-1 and 5 CSIs (3 Senior, 2 Volume) in Scenario-2, for whom the eye tracking data were of sufficient quality to be incorporated in the final analysis.



Figure 6.3. Snapshot of a participant's point of view in the study, illustrating a fixation on the mobile phone on the floor (intersection of the blue lines).

6.2.8.1. Eye Movement Characteristics

Cumulative Dwell Time Percentages on the AOIs

After consideration of the total viewing times, dwell times on AOIs are considered. The first comparison is between fixations on Placed AOIs (Figure 6.4a). The results revealed that scenario did not affect the time spent on Placed AOIs (Placed AOIs, M =55.21, SD=13.88 (Scenario-1); M =60.58, SD=13.36 (Scenario-2); $t(9)=-0.65$, $p=0.53$, $d=-0.39$)

When comparing the two groups of CSI, dwell times on Placed AOIs were found to be slightly longer for the less experienced volume CSIs, in line with my predictions. However, this difference with senior CSIs was not statistically significant, possibly due to the low

numbers of participants in the study (Placed AOIs, $M = 64.53$, $SD = 15.46$ (Volume); $M = 51.92$, $SD = 8.56$ (Senior); $t(9) = 1.72$, $p = 0.12$, $d = 1.01$).

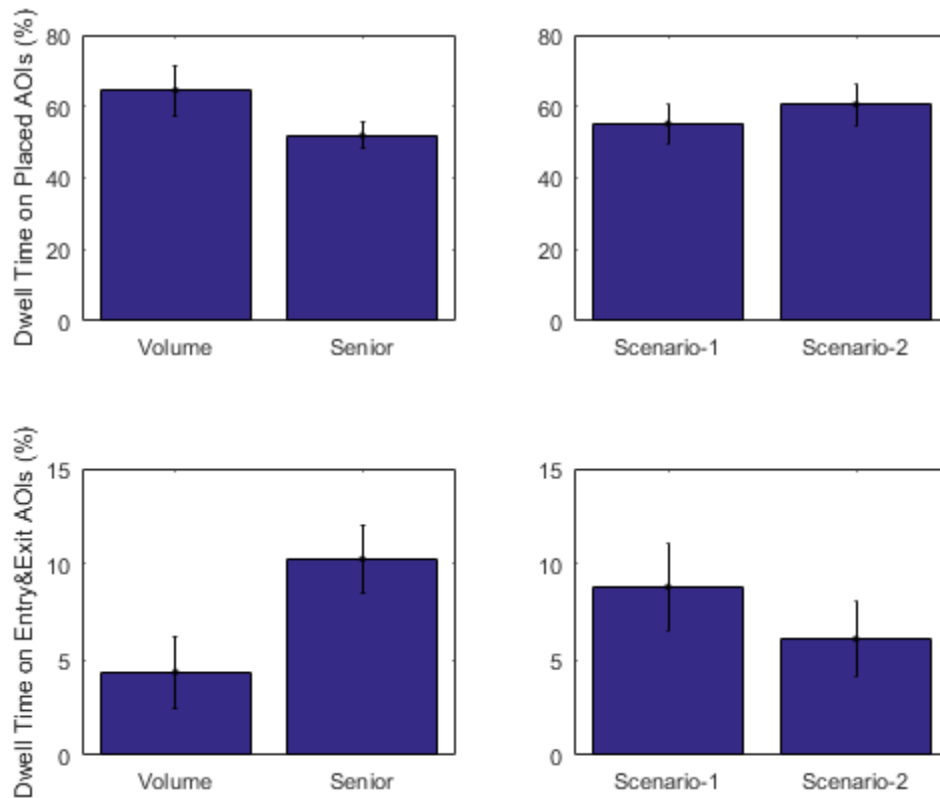


Figure 6.4. a) Mean cumulative dwell times on Placed AOIs. b) Mean cumulative dwell times as a percentage of the overall viewing times on Entry and Exit Points. On the right, the two scenarios are compared, while on the left subplot compares the two CSI groups. Error bars in the subplots show the standard error of the mean across participants ($N = 11$).

Viewing behaviour towards “Points of Entry or Exits” (see figure 6.4b) was found to be influenced significantly by the level of experience. Senior CSIs were found to spend significantly longer looking at Entry and Exit points than less experienced volume inspectors ($M = 4.35$, $SD = 4.16$ (Volume CSIs); $M = 10.26$, $SD = 4.23$ (Senior); $t(9) = -2.33$, $p = 0.045$, $d = -1.41$). Scenario, in contrast, did not influence dwell times on Entry and Exit points ($M = 8.81$, $SD = 5.64$ (Scenario-1); $M = 6.09$, $SD = 4.33$ (Senior); $t(9) = 0.88$, $p = 0.4$, $d = 0.54$).

A comparison of the dwell times on non-classified objects (uncategorized AOIs, see Figure 6.5a) showed that even though senior CSIs spent more time exploring non-classified objects than volume CSIs, the difference was not statistically significant ($M = 31.12$, $SD = 12.06$ (Volume); $M = 37.81$, $SD = 5.82$ (Senior); $t(5.54) = -1.14$, $p = 0.3$, $d = -0.71$). Scenario did not have an effect on these viewing times either ($M = 35.98$, $SD = 10.1$ (Scenario-1); $M = 33.33$, $SD = 9.28$ (Scenario-2); $t(9) = 0.45$, $p = 0.66$, $d = 0.27$)

Finally, dwell times on “Possible Evidence” AOIs (which were described in the verbal reports) were compared (see Figure 6.5b). Again, scenario did not have an effect on the time spent on “Possible Evidence” ($M = 39.8$, $SD = 9.89$ (Scenario-1); $M = 46.48$, $SD = 12.72$ (Scenario-2); $t(9) = -0.98$, $p = 0.35$, $d = -0.59$). Volume CSIs had a tendency to allocate more attention to “Possible Evidence” in the scene than senior CSIs, however the difference was not statistically significant ($M = 47$, $SD = 13.32$ (Volume); $M = 39.37$, $SD = 8.84$ (Senior); $t(9) = 1.14$, $p = 0.28$, $d = 0.68$).

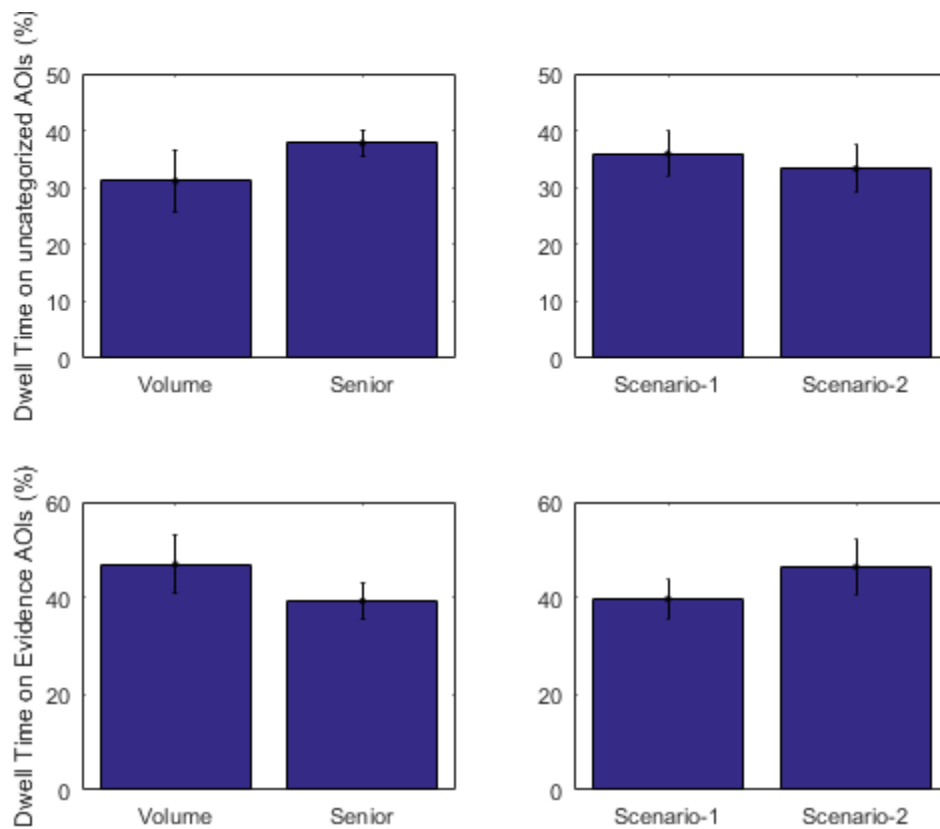


Figure 6.5. a) Mean cumulative dwell times on Uncategorized AOIs. b) Mean cumulative dwell times (as a percentage of the overall viewing times) on possible evidence AOIs (i.e., items mentioned in the verbal reports). On the right, the two scenarios are compared, while on the left subplot compares the two CSI groups. Error bars in the subplots show the standard error of the mean across participants (N=11).

The comparisons above pool data across the other dimension, before analysis. The next analyses will examine how expertise influenced viewing times within each scenario. For both scenarios, it was found that volume CSIs spent more time looking at “Possible Evidence” than senior CSIs, but this difference was not statistically significant (Scenario-1; Placed AOIs, $M = 57.98$, $SD = 17.27$ (Volume); $M = 52.43$, $SD = 12.66$ (Senior); $t(4) = 0.45$, $p = 0.68$, $d = 0.37$; Evidence AOIs, $M = 41.31$, $SD = 10.17$ (Volume); $M = 38.29$, $SD = 11.59$ (Senior); $t(4) = 0.34$, $p = 0.75$, $d = 0.28$; Scenario-2; Evidence AOIs, $M = 55.54$, $SD = 16.11$ (Volume); M

=40.45, SD=7.58 (Senior); $t(3)=1.48$, $p=0.24$, $d=1.2$). There was one exception: For scenario 2, volume investigators spent more time looking at Placed AOIs (Scenario-2; Placed AOIs, $M=74.34$, $SD=6.26$ (Volume); $M=51.41$, $SD=4.68$ (Senior); $t(3)=4.78$, $p=0.017$, $d=4.15$).

Senior CSIs allocated more attention to the Entry and Exit Points and the rest of the scene (Uncategorized assets) than volume CSIs for both scenarios (Scenario-1; Entry and Exit Points, $M=6.3$, $SD=4.5$ (Volume) ; $M=11.32$, $SD=6.36$ (Senior); $t(4)=-1.12$, $p=0.033$, $d=-0.91$; Uncategorized, $M=35.71$, $SD=13.86$ (Volume) ; $M=36.24$, $SD=7.93$ (Senior); $t(4)=-0.06$, $p=0.96$, $d=-0.05$; Scenario-2; Entry and Exit Points, $M=1.42$, $SD=0.78$ (Volume) ; $M=9.21$, $SD=1.02$ (Senior); $t(3)=-10.29$, $p=0.002$, $d=-8.58$; Uncategorized, $M=24.25$, $SD=6.34$ (Volume) ; $M=39.38$, $SD=3.81$ (Senior); $t(3)=3.45$, $p=0.041$, $d=-2.89$). Note that these comparisons are based on a small sample size, which also differed across the groups. Before drawing strong conclusions on the basis of this data, it is therefore important that the results are replicated using larger and more balanced samples.

The above data suggest that experience may have an effect on viewing patterns in crime scene investigations. To investigate this role of experience in more detail, the effects of years of experience was determined (a continuous measure, possibly less influenced by other factors than those determining who gets promoted to senior CSI). Figure 6.6 provides a scatterplot of the association between years of experience and viewing time on Placed AOIs, revealing a significant negative correlation ($r=-0.63$, $p=0.038$), indicating that inspectors with more years of experience fixate a “Placed” AOIs less. A similar analysis for possible Entry and Exit AOIs showed a similar trend towards a correlation between years of experience and cumulative viewing times ($r=0.54$, $p=0.084$).

Correlation between the exploration time for the uncategorized rest of the scene and the CSI experience was in line with the correlation observed for points of Entry and Exit and was marginally significant ($r=0.6$, $p=0.05$). No correlation observed between the time percentages on “Evidence AOIs” but the correlation was also negative, ($r=-0.38$, $p=0.25$).

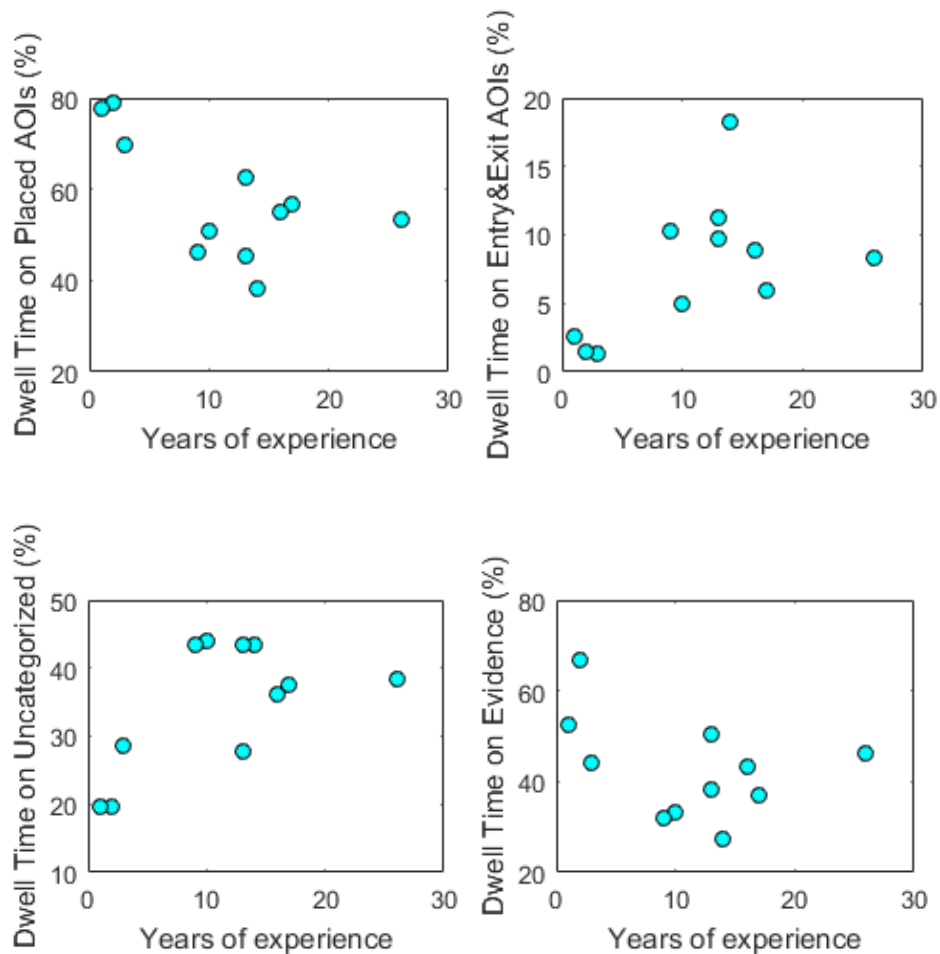


Figure 6.6. Scatter plots showing the association between years of experience and dwell times on AOIs

Inspection of the scatter plots in Figure 6.6 suggests that the three least experienced investigators have a different viewing pattern from the other investigators. To test this observation statistically, I decided to compare the investigators with 1-3 years of experience

to the senior CSIs. In order to have equal sample sizes, senior CSIs were divided into two groups with more experience (14-26 years) and less experience (9-13 years). This analysis was performed across the two scenarios.

Viewing times on Placed AOIs differed significantly across the three groups with longer dwell times for inexperienced volume CSIs than either senior CSI group (Placed AOIs, $M = 75.45$, $SD = 4.83$ (Inexperienced Volume) ; $M = 48.95$, $SD = 12.39$ (Less Experienced Senior); $t(4) = 3.45$, $p = 0.026$, $d = 2.82$; $M = 54.9$, $SD = 1.7$ (Experienced Senior); $t(4) = 6.96$, $p = 0.002$, $d = 5.68$), while no significant differences were found in the viewing times of Evidence AOIs between the different CSI groups (Evidence AOIs, $M = 54.58$, $SD = 11.51$ (Inexperienced Volume) ; $M = 36.51$, $SD = 12.24$ (Less Experienced Senior); $t(4) = 1.86$, $p = 0.14$, $d = 1.52$; $M = 42.23$, $SD = 4.56$ (Experienced Senior); $t(4) = 1.73$, $p = 0.16$, $d = 1.41$),

There was a significant correlation between years of experience by each CSI and cumulative time percentage on “Placed AOIs” ($r = -0.84$, $p = 0.038$). A similar correlation found more experienced senior CSIs and inexperienced volume CSIs ($r = -0.94$, $p = 0.005$).

However, the viewing strategies on the Entry and Exit points were shorter for the inexperienced volume CSIs than either senior groups ($M = 1.82$, $SD = 0.7$ (Inexperienced Volume) ; $M = 12.79$, $SD = 4.8$ (Less Experienced Senior); $t(2.08) = -3.92$, $p = 0.055$, $d = -3.2$; $M = 7.74$, $SD = 1.61$ (Experienced Senior); $t(4) = -5.85$, $p = 0.004$, $d = -4.77$) Also, dwell times showed a similar tendency while viewing the rest of the scene ($M = 22.74$, $SD = 5.19$ (Inexperienced Volume) ; $M = 38.26$, $SD = 9.1$ (Less Experienced Senior); $t(4) = -2.57$, $p = 0.062$, $d = -2.1$; $M = 37.37$, $SD = 1.2$ (Experienced Senior); $t(2.21) = -4.76$, $p = 0.034$, $d = -3.88$)

In addition, analysis of years of experience and dwell times on Entry and Exit points and the rest of the scene (the Uncategorized objects and places) were revealed similar correlations between inexperienced volume CSIs and each of the senior CSI groups. (The correlation between inexperienced volume CSIs and less experienced senior CSIs; for Entry and Exit points, $r=0.92$, $p=0.01$, for the Uncategorized $r=0.73$, $p=0.1$; between inexperienced volume CSIs and experienced senior CSIs; for Entry and Exit points, $r=0.91$, $p=0.013$, for Uncategorized $r=0.91$, $p=0.012$)

Within Experience Groups Analysis on each AOI

Previous analyzes did not suggest an effect of scenario, but this may have been due to the pooling of experience levels. A new set of analyzes therefore testing the effect of scenario within each of the participant groups. As before, no effects of scenario were found, further confirming that scenario did not influence viewing times within each participant groups. (Within Senior CSIs; Total Duration Time, $M=433.41$ $SD=257.9$ (Scenario-1), $M=302.04$ $SD=100.7$ (Scenario-2), $t(4)= 0.82$, $p=0.46$, $d=0.67$; Placed AOIs, $M= 52.43$ $SD=12.66$ (Scenario-1), $M=51.41$ $SD=4.67$ (Scenario-2), $t(4)= 0.57$, $p=0.6$, $d=0.11$; Entry&Exit AOIs, $M=11.32$ $SD=6.36$ (Scenario-1), $M=9.21$ $SD=1.02$ (Scenario-2), $t(4)= 0.13$, $p=0.9$, $d=0.46$; Uncategorized, $M=36.24$ $SD=7.93$ (Scenario-1), $M=39.38$ $SD=3.81$ (Scenario-2), $t(4)=-0.62$, $p=0.57$, $d=-0.5$; Evidence AOIs, $M=38.29$ $SD=11.59$ (Scenario-1), $M=40.45$ $SD=7.57$ (Scenario-2), $t(4)= -0.27$, $p=0.8$, $d=-0.22$; Within Volume CSIs; Total Duration Time, $M=355.97$ $SD=206.78$ (Scenario-1), $M=382.78$ $SD=103.6$ (Scenario-2), $t(3)= -0.16$, $p=0.88$, $d=-0.16$; Placed AOIs, $M=57.98$ $SD=17.26$ (Scenario-1), $M=74.34$ $SD=6.26$ (Scenario-2), $t(3)= -1.23$, $p=0.31$, $d=-1.26$; Entry& Exit AOIs, $M=6.3$ $SD=4.5$ (Scenario-1), $M=1.41$ $SD=0.8$ (Scenario-2), $t(3)= 1.46$, $p=0.24$, $d=1.51$; Uncategorized, $M=35.71$ $SD=13.85$ (Scenario-1), $M=24.25$ $SD=6.34$ (Scenario-2), $t(3)= 1.06$, $p=0.37$, $d=1.06$; Evidence AOIs,

M=41.31 SD=10.16 (Scenario-1), M=55.54 SD=16.11 (Scenario-2), $t(3) = -1.25$, $p=0.3$, $d=-1.05$)

CRQA of Eye Movement Data

The next step of the analysis involved a CRQA of the viewing patterns to examine whether the overlap in viewing patterns was higher within a CSI group (volume versus senior) than between groups. This was achieved by comparing eye movement patterns of pairs of CSIs and then pooling the numbers quantifying the overlap for pairs within a CSI group (volume or senior) and pairs between groups (comparing volume with senior CSIs). This resulted in three types of comparisons: (1) If the two participants compared were both volume CSIs, this was a V-V comparison, (2) if both were senior CSIs, this was an S-S comparison, and (3) if one was volume CSI and another a senior CSI, this was a V-S comparison. These were then pooled into a V-V and S-S average and a V-S average, which were then compared statistically using one way ANOVAs. Because a Shapiro-Wilk's test on Recurrence suggested significant deviations from the normal-distribution assumption, square root transformations were applied to the data. Also, because of the unequal group sizes (V-V group 10 comparisons, S-S group 15 comparisons, V-S group 30 comparisons), there may be issues with the Homogeneity of Variance assumption. To counteract such issues, non-parametric comparisons were also conducted between the groups.

The means were shown in Figure 6.7. The results revealed no effect of experience on CRQA measures irrespective of scenario differences. All groups showed very similar perceptual patterns for recurrence percentages or determinism in the investigation process. (Recurrence; $F(2,52)=0.35$, $p=0.72$, $\eta^2 =0.01$; Determinism; $F(2,52)=0.49$, $p=0.61$, $\eta^2 =0.02$). Laminarity scores were low for some comparisons, meaning that no comparison could be made on this

measure. These results, however, may have been affected by the loss of eye tracking data and should be interpreted carefully. (Kruskall Wallis non parametric tests, Recurrence; $H(2)=0.48$, $p=0.79$, Determinism $H(2)=1.03$, $p=0.6$)

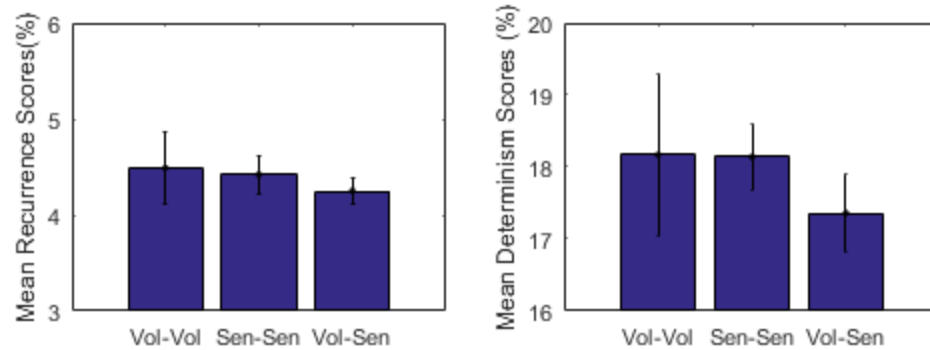


Figure 6.7. a-b) Between group and within group average CRQA measures: Recurrence, Determinism. Vol=Volume CSIs, Sen=Senior CSIs. Error bars show the standard error of the mean across participants.

Total Duration Time at the Scene

Mean total duration times are shown in Figure 6.8 across scenarios and participant groups. Total viewing time was found to be independent of the scenario. There was a trend towards longer inspection times for Scenario-1, but this difference was not statistically significant ($M=394.69$ sec, $SD=213.33$ (Scenario-1), $M=334.33$ sec, $SD=98.54$ (Scenario-2); $t(7.3)=0.62$, $p=0.55$, $d=0.36$).

The results also showed that Volume and senior CSIs had similar inspection times ($M=366.69$ sec, $SD=155.82$ (Volume CSIs); $M=367.73$ sec, $SD=189.31$ (Senior CSIs); $t(9)=-0.1$, $p=0.99$, $d=-0.006$).

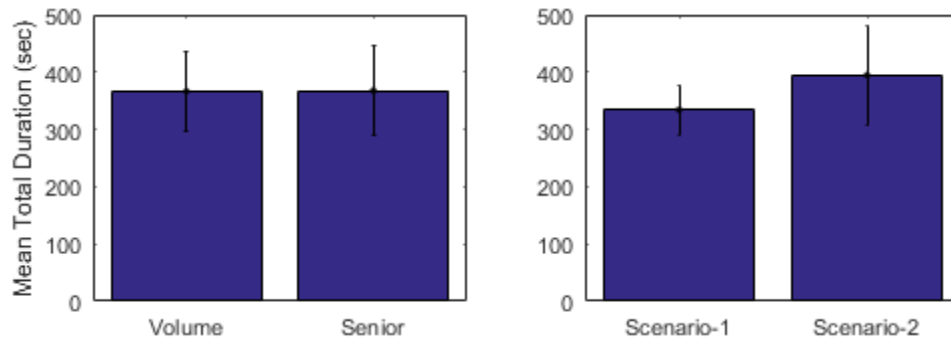


Figure 6.8. Total viewing time. comparing the differences between scenarios (right plot) and the two levels of CSI experience (left plot). Error bars in the subplots show the standard error of the mean across participants (N=11).

6.2.8.2. Verbal Protocols

The eye movement patterns suggest that volume CSIs spend more time on Placed AOIs than senior CSIs. The question therefore arises whether verbal reports show a similar trend, with volume CSIs reporting more about these objects compared to senior CSIs. Therefore, the analysis was conducted while comparing verbal reports from volume and senior CSIs irrespective of the scenario difference.

The coding process is illustrated in the following example:

(CSI-3 reporting in Scenario-2) “.....Looking at the knife thinking, DNA. It’s a bit textured for fingerprinting but that would be recovered. On the floor as well there was some staining that looked like potentially, it could have been something that had spilt or been cleaned up. And on the door, near the handle as well, there was a little bit of what looked like trace but it could be tea stains, but I wasn’t sure, so that would be something else I’d be testing, potentially for blood. Yes, I’ve actually pointed that, and I go and look over that and I think,

yes, that's clearly a red herring because it's a piano. And then that kind of throw there, I don't know if that was part of it or if it's just how the room normally is. Has somebody been lying or sleeping on the sofa maybe, an extra person?"

Table 6.3.

An example of coding process.

<i>Coding Scheme</i>	
<i>Referring to object</i>	<ol style="list-style-type: none"> 1. Knife 2. some staining 3. there was a little bit of what looked like trace but it could be tea stains
<i>Referring to evidence</i>	<ol style="list-style-type: none"> 1. Thinking DNA 2. a bit textured for fingerprinting but that would be recovered 3. so that would be something else I'd be testing
<i>Referring to Activities(Inferences)</i>	<ol style="list-style-type: none"> 1. something that had spilt or been cleaned up 2. Has somebody been lying or sleeping on the sofa maybe, an extra person?
<i>Referring to Scene Features</i>	<ol style="list-style-type: none"> 1. on the floor 2. and on the door, 3. near the handle as well 4. because it's a piano 5. that kind of throw (on the sofa) there

Two coders were involved in the coding the verbal data. One coder coded 100% of the verbal reports, while a second coder coded 25% of the verbal data segments independently. To assess the correspondence between the coding of the two coders, the test inter-rater reliability Mezzich's K statistic was computed for the overlapping data. This revealed a value of $\kappa = 0.67$, indicating substantial agreement between the coders (Landis & Koch, 1977).

Overall, statements were often about objects and the scene. Interestingly, statements about evidence (interpretation of the objects) were few. While conducting a 2 (Experience) x 4 (Type of References) ANOVA because standardized residuals suggested significant deviations from the normal-distribution assumption in Shapiro-Wilk's test and logarithmic transformations were applied to the data and further analysis was conducted for the comparisons between the groups. The untransformed data was plotted for easy interpretation of the reported results (see Figure 6.9.)

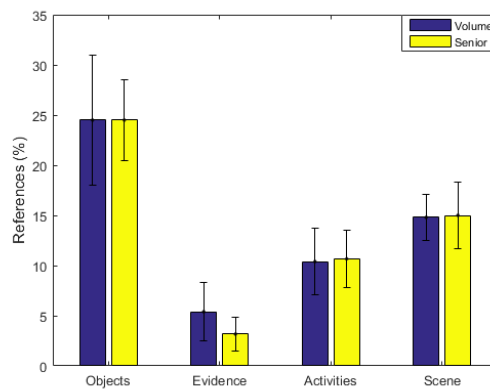


Figure 6.9. References in the verbal reports to objects, evidence, the scene and activities by the two groups of CSIs.

Experience did not have an effect on references stated in the verbal reports ($F(1,9)=0.14$, $p=0.72$, $\eta^2 = 0.02$). In contrast, an effect of type of references was found ($F(3,27)=28.44$, $p<0.001$, $\eta^2 = 0.09$). in the absence of an interaction effect between experience and type of references ($F(3,27)=0.88$, $p=0.47$, $\eta^2 = 0.09$) meaning that experience did not have an influence on the references stated in the verbal reports.

Also further analysis revealed that in contrast to the pattern in the eye movement data, the verbal reports do not suggest a stronger focus of the less experienced volume CSIs on the evidence, or a stronger focus on the scene from the more experience senior CSIs (Reference to Objects; $t(9)=-0.28$, $p=0.79$, $d=-0.16$, reference to Evidence; $t(9)=1.01$, $p=0.34$, $d=0.6$; reference to Activities $t(9)=-0.05$, $p=0.96$, $d=-0.03$; Scene Features; $t(9)=0.16$, $p=0.88$, $d=0.1$).

6.2.8.3. Prioritized Evidence

To examine whether the scenario influenced mentioning and prioritizing of evidence, the number of CSIs that reported one of 12 different objects that could present possible evidence were compared across scenarios (Figure 6.10). The number of times some of the evidence was mentioned (as prioritized) appeared to differ across scenarios. For example, blood was more often referred to for Scenario 1 than for Scenario 2. Conversely, the glass and the bleach were more often mentioned for Scenario 2. These findings suggest that verbal prioritization and selection of evidence is dependent on the context of the crime scene. Numbers, however, were small, and therefore did not allow for statistical comparison (the assumptions of the appropriate statistical tests would be violated strongly).

The plots suggest that the evidence most referred to was the blood on the floor (9 out of 11 (82%) participants considered it as important and relevant to the case). The least important evidence was the milk: Only one participant (9%) reported this as relevant.

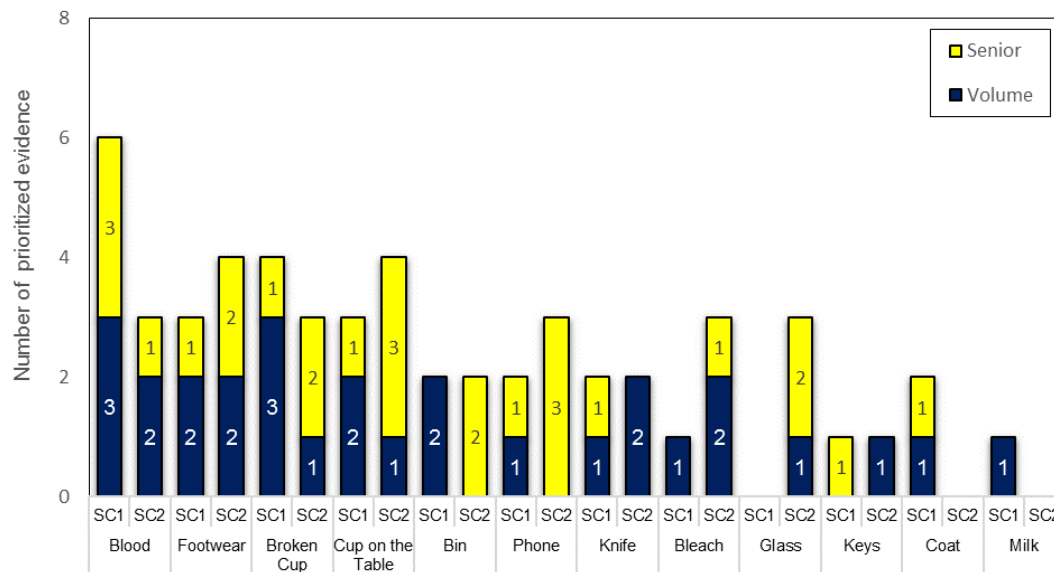


Figure 6.10. Selected and Prioritized Evidence (yellow and blue stacked bars) across participant groups and scenarios (N=11).

6.2.8.4. Case Hypothesis – Role of scenario

Scenarios appeared to influence the evidence considered in the verbal reports, but not in the eye movement data. The question arises whether scenario and CSI group influenced the hypotheses about the case. Analysis of the verbal reports suggested three main hypotheses about the case (see Figure 6.11).

Participants generally generated only one hypothesis, with the exception of a senior CSI who generated two hypotheses. 3 out of 6 participants in Scenario-1 also suggested self-harm (and suicide) as a possible hypothesis. Whether a crime was hypothesized appeared to depend on the scenario. 3 out of 5 participants in Scenario-2 did not find enough evidence to conclude a crime had been committed. Overall, 6 out of 11 participants suspected an assault (e.g. domestic incident, a light injury, an occurrence with somebody).

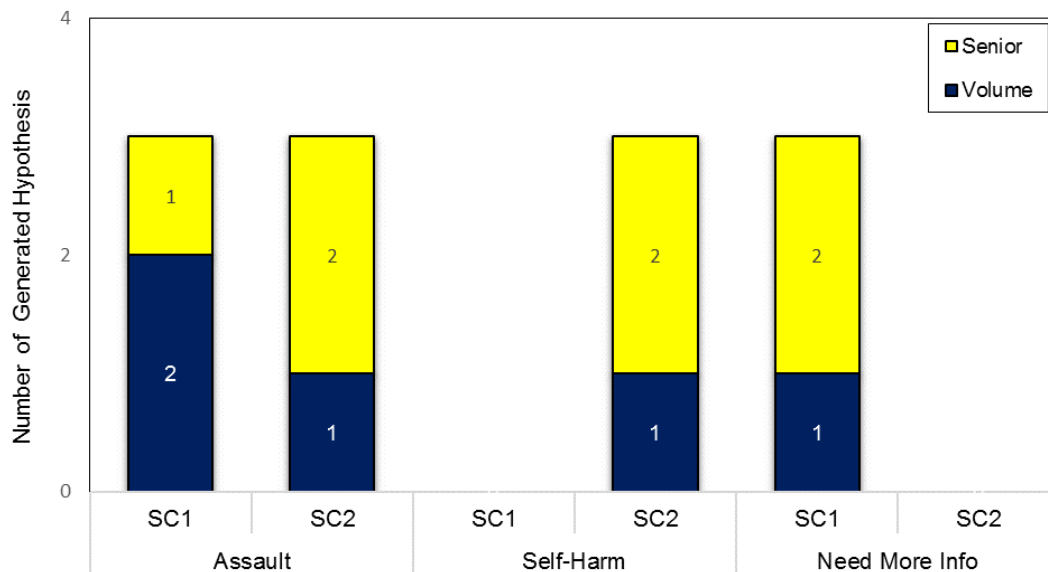


Figure 6.11. Generated Hypothesis for participant groups on each scenario (N=11).

6.3. Discussion

To complement the static image, desk-mounted eye tracking studies in Experiments 1 and 2, Experiment 3 used mobile eye tracking during actual inspection of a (simulated crime scene). Eye movements, inspection times, and verbal reports were compared across volume (less experienced) and senior (more experienced) CSIs across two scenarios (although the actual scene itself was identical across scenarios). Results suggest no differences in inspection times across CSI groups or scenarios. Eye movements suggested that the more experienced senior CSIs spent more time looking at the broad layout of the crime scene (floors, ceilings, exit and entry points), while the less experienced CSIs spent more time looking at Placed objects and possible evidence. Verbal reports, in contrast, did not reveal group differences, but did suggest an effect of scenario. This suggests that eye movements measure different aspects of the crime scene investigation process than verbal reports, and may be more relevant when attempting to assess the experience level of a CSI than verbal reports.

Time spent to complete the task was used as a performance indicator, as past literature had suggested this as a valid measure of expertise (Gegenfurtner et al., 2011, for an overview; Baber & Butler, 2012, investigating volume crime investigation). In contrast to these previous reports, no difference in terms of time spent at the scene was found between volume and senior CSIs, casting doubt on the relevance of inspection time as an indicator of expertise. This result was unexpected as, for example, Baber and Butler (2012) observed that experts were faster than novices. Likewise, Schraagen and Leijenhurst (2001) found more efficient visual searching strategies of experts in time-constrained tasks, and also found that extending the time available merely led to exploration of more evidence. The lack of an expertise effect on inspection times in the present study is in line with Santtila, Korpela and Hakkanen (2004), who found no effect of expertise on investigation of linking car crimes.

The question lies in the difference amongst studies. One reason could be the experience levels used. Baber and Butler (2012) and Schraagen and Leijenhurst (2001) compared two levels of experience. In contrast Santtila et al. (2004) compared four levels of experience. The novices in Baber and Butler's (2012) study were forensic students, while novices in Schraagen and Leijenhurst's (2001) study were inexperienced interns. Their novices therefore lacked experience with the crime scene investigation process. In contrast, novices in Santtila et al.'s (2004) study were inexperienced detectives already working in a police force, and therefore had some experience with crime investigation. Novices in the present study were less experienced CSI, having experience with crime scene investigation. Apart from training, the main difference in the experts and novices in the present study lies mostly in the experience with serious cases (rape, murder etc.). Volume CSIs mainly work on cases classified as non-serious such as burglary, criminal damage to properties. In serious cases (e.g. murder, rape), volume CSIs do not have a big role, but they are expected to assist senior

CSIs. Together, the evidence suggests that experience with some form of crime scene investigation in the novice group can explain the lack of difference in inspection time across experts and novices. The confusion about what consists a novice is not unique to studies of crime scene investigation. Also in studies of surgery, novices vary broadly in their experience (for an overview, see Hermens et al., 2013). More studies using varying levels of ‘novices’ are therefore needed to out to validate whether time spent to complete the task is a reliable indicator of experience in crime scene process.

Whereas overall time spent at the crime scene was unaffected by experience, eye movement patterns did show a difference across the two groups of CSIs. These results were further supported by a correlation with years of experience and eye movement patterns. As the years of experience on the job increased, the time spent on “Placed AOIs” decreased across the participants. In contrast, attention to “Entry and Exit Points” and other parts of the scene correlated positively with years of experience. One interpretation of the data is that in the beginning of volume CSIs’ career, a significant part of the crime scene investigation focuses on collecting evidence, while a broader approach towards approaching a crime scene only arises with experience. Such an interpretation is in line with Baber and Butler (2012), who suggested that less experienced CSIs used a stronger bottom-up approach, contrasting the more top-down approach of more experience CSIs. For less experienced CSIs, a crime scene may seem more full of conflicting objects and materials to be considered. They may therefore concentrate more on the objects that are “out of place” (which were placed deliberately) in the scene. The lack of attention of the less experienced CSIs to “Entry and Exit” points (the overall scene) suggest a relatively constrained cognitive approach while assessing the scene.

The finding of a stronger focus towards exit and entry points in experts is supported by the findings of Baber (2010), who defined the process of starting the case from entry and exit point as a sign of “intuition” in CSIs, possibly linked to previous experience and domain related knowledge. The longer dwell times on wider parts of the scene by experienced senior CSIs agrees with the concept of ‘situation awareness’ (Endsley, 2000), which has been linked to expertise (Koh et al.,2011).

While group differences were found in the AOIs analysis, the analysis of the patterns of eye movements (using Cross-Recurrence Analysis) did not reveal an effect of experience. This lack of an effect may be due to the large variability (large SDs) across CSIs within a group, reflecting diverse visual search strategies (recurrence measures were low overall). In the analysis, the entire inspection interval was compared. It may be possible that parts of the inspection show higher agreement across CSIs (e.g., the first 20 seconds), but it is unclear how to identify such intervals with higher overlap objectively.

While eye movements were affected by expertise, no such expertise effects were found on the verbal reports. Instead, verbal reports suggested a role of the scenario provided. A possible explanation of such an effect of scenario would be that it influenced the very initial stages of the crime scene investigation process, by influencing the initial hypothesis. For example, in the “lost girl” scenario, verbal reports suggested that participants were more likely to refer to self-harm incident, even while objects in the scene indicated a possible struggle and another person’s presence in the scene. These results illustrated the importance of expectations in overall crime scene examination process (Dror, 2015). Such expectation effects agree with other studies indicating the importance of advance information (Ballard &

Hayhoe, 2009; Land, 2009; Tatler, Hayhoe, Land & Ballard, 2011; Orquin & Mueller Loose, 2013)

The results suggest a role of years of experience, possibly mediated by the number of cases inspected, in how a crime scene is approached. For example, Ericsson (2006) suggested that around 10 years of experience are necessary to achieve good performance. The present eye movement data suggest that years of experience and continuous engagement in deliberate practices from diverse scenes may lead the CSIs to a more holistic and thorough information processing and reasoning. Eye movements data also suggest that volume CSIs mostly rely on standard procedures and guidelines, as reported earlier. Such an interpretation is in line with work by Chi, Feltovich and Glaser (1981) who found that experts and novices mostly differ in how they organize and adapt their domain knowledge, rather than in their problem solving ability.

The data are suggestive of the use of various strategies by the CSIs, in line with research in medical domains (see also Patel et al., 2012). One of such strategies is to make use of similarity to previous cases. Baber (2010) describes this similarity based approach as a process guided by “typical situations” based on previous experience and knowledge. This is illustrated by the following quote from the verbal report of an experienced CSI:

“...And then there’s always a bin, you’ve got to look in the bins, haven’t you? There’s always the bin to route through. Inside the bin... I always learnt early in career to look at the ceilings, any splatter pattern or anything else, people tend to look down and not look up. And then...”

The study was not without limitations. The main limitation was probably the size of the sample, which is always an issue when dealing with highly skilled experts. There are not many experts in the first place (which is why they are experts) and as there is high demand for them, they are generally too busy to take part in studies (c.f., the case with consultant surgeons; Hermens et al., 2013, for a discussion). A related issue is that participants were all recruited from a single Police department. Reasons were the lack of cooperation of other departments, and limited resources to conduct the study, preventing a broader, multi-site investigation. A further limitation was the use of a single case and a single site. This relates to the limited availability of experts, so that it was not possible to test experts across multiple sites due to time constraints. Another reason was the use of simulated rather than an actual crime, which meant that the site needed to be kept in place and undisturbed across the entire testing interval. The way this was achieved was by having the simulated crime scene inside, which may have limited the results to indoor settings only. As discussed before, the expertise comparison was limited to a comparison of volume and senior comparisons, who were both relatively experienced, which may have limited the expertise effects. This was again, due to practical reasons. In order to be able to recruit sufficient experts, the simulated crime scene was placed close to the police headquarters, which made it difficult to bring in other study groups, such as forensic students. While the eye tracker could be easily moved towards the participants, this was not the case for the crime scene. This is where studies using static images presented on a computer screen have an advantage, as it is possible to bring both the eye tracker and the scene to the participants.

Despite the various limitations, the study had several strengths. The main strength of the study was the actual exploration of a crime scene and the use of state-of-the-art mobile eye tracking technology to not only assess the overall time spent at the scene and the verbal

reports of the participants, but also the visual exploration of the scene. Using a simulated crime scene meant that the task performed by the participants closely resembled their day-to-day activities, in a much stronger way than would be possible using static images of crime scenes on a computer screen. The mobile eye tracking data revealed differences between experts and novices that were not revealed by the measures that would have been obtained without eye tracking technology (time spent, and verbal reports). The eye tracker also enabled the retrospective verbal reports, by providing participants an accurate account of their scanning behaviour (recorded video images and superimposed eye gaze cursor), which is likely to have led to more accurate verbal reports than would have been possible without such data.

While mobile eye tracking adds ecological validity to the study, the quality of the mobile eye tracking data was lower than for desk-mounted eye tracking. Lighting conditions vary when participants walk around and explore the surrounding, the head-band of the eye tracker may move during exploration, and participants may look outside the field of view of the eye tracker (e.g., when looking down using the eyes only), all provide challenges to accurate eye tracking (Holmqvist et al. 2011). Current versions of mobile eye trackers are limited in their sampling rate, both of the video image and the eye tracking data. Attempts to increase these sampling rates will need to be accompanied with more efficient storage and processing of the data, which will possibly require new development in computer technology. Also, one major drawback of the use of mobile eye tracking, is that it is extremely time-consuming to analyze the data and, in part, subjective (in the use of the library for coding the data, and when interpreting eye movement data influenced by measurement noise). It is not until computer vision achieves sufficient standards to automatically detect and classify objects, that this issue will be alleviated. While participants can be expected to mostly use overt

(moving their eyes) shifts of attention, it cannot be excluded that they also used covert (no eye gaze shift) shifts of attention, and therefore eye movements remain a somewhat restricted way of uncovering attention (Foulsham, Chapman, Nasiopoulos & Kingstone, 2014). Despite this range of possible issues, the large amount of data collected is likely to filter out many of the problems in the average data. Participants whose data were of poor quality were excluded from the analysis, and coding of the data was discussed with another person when there was uncertainty. Calibration at the start and the end of the crime scene investigation process allowed for an assessment of the amount of head band slippage. Practice with walking around with the eye tracker before the start of the crime scene investigation ensured that participants got used to wearing the eye tracker. While for the study of saccade characteristics, higher sampling rates are required, the sampling rate of the mobile eye tracker was sufficient for the present purpose of the study: To evaluate the point of gaze during fixations.

6.4. Conclusion

Mobile eye tracking during actual (simulated) crime scene investigation reveals differences between less experienced and more experienced CSIs not revealed by time spent in the scene and verbal reports. These results suggest an important role of eye tracking in the assessment of expertise in crime scene investigation, which needs to be explored in future studies using larger and more diverse samples and a broader range of crime scene.

Chapter 7. Effects of Training on Behavioural and Evaluative Responses during Viewing Emotional Stimuli

7.1. Introduction

Current research has largely neglected the role of perceptual and cognitive processes in forensic sciences (Dror, 2015). Instead, technological advances have attracted the bulk of attention (Kelty et al., 2011). While the importance of human factors has been recognized in other disciplines (Mearns, Flin, Gordon & Fleming, 2001; Reader, Flin, Lauche & Cuthbertson, 2006), research on cognitive factors in forensic science only recently emerged (Dror & Cole, 2010).

One aspect of cognition involved in forensic sciences, is how people respond emotionally to crime related stimuli. Studies have suggested a strong link between emotion and cognition and implicated shared brain structures such as the amygdala (Phelps, 2006). Given this strong link between cognition and emotion, and given the nature of the crime scenes, it is therefore important to study how the emotional aspects of the crime scene investigation process influences cognition during the work.

Work on emotion in crime scene investigation has largely focused on the physiological responses underlying the stress response or the role of emotion for the well-being of CSIs and to prevent burnout (Kelty & Gordon, 2015; Schaible & Gecas, 2010; Sewell, 1994). For example, studies of heart rates during crime scene investigations have demonstrated that not only homicide case, but less severe cases, such as burglary, lead to increased heart rates.

Increased experience was found to reduce these effects (Adderley, Smith, Bond & Smith, 2012). Others have examined the symptoms of stress during crime scene investigations, and the coping mechanisms to deal with this stress (Miller, 2008). Likewise, a further study examined the role of humour in dealing with crime scene investigation stress (Vivona, 2014).

It is important to distinguish between the overall emotional state of the CSI and the immediate emotional effects of the stimuli in the crime scene. The above studies are mostly concerned with the former aspect of emotion. Studies dealing with the latter aspect often involve studying the response to emotional stimuli. These latter studies, for example, have shown that when shown with pairs of positive, negative, and neutral images, participants tend to focus on emotional stimuli (Carniglia, Caputi, Manfredi, Zambarbieri & Pessa, 2012). Thematic content of the images, however, was also found to play a role. A similar result was obtained by Humphreys, Underwood and Chapman (2010), who found that emotionally positive stimuli received more attention, but did not lead to better memory of the contents of the stimuli. In contrast, Chipchase and Chapman (2013) found that emotionally negative stimuli attracted more attention than positive stimuli. These effects may be modulated by age of the observer, which can be an important aspect when comparing expert and novice CSIs. The elderly show tendency to focus on positive images and ignore negative images (Reed, Chan & Mikels, 2014).

While these studies examined what type of stimuli attract more attention, other studies have tried to establish what parts of emotional stimuli attract the observer's eye (e.g., Bradley, Houbova, Miccoli, Costa & Lang, 2011; Lanata, Valenza & Scilingo, 2012). For example, Bradley et al. (2011) found that more fixations and larger amplitude saccades were made for emotionally loaded (positive and negative) images. Similar results were obtained by

Carniglia et al. (2012), who found longer fixation durations and more fixations for emotional stimuli.

The above mentioned studies focused on the influence of emotion on attention. Another aspect of cognition is memory and the same studies have also shown that emotional contexts have a tendency to leave stronger memory traces than neutral contexts. For example, it has been shown that enhanced memory in younger adults, when compared to older adults, is restricted to negative images (Kensinger, Garoff-Eaton, & Schacter, 2007) and not all details of the contexts are remembered, however the central details are recalled better. When presented in peripheral vision, memory is reduced for detail when emotional stimuli are used (Chipchase & Chipman, 2013). The finding that people zoom in on a weapon and have poor memory for what else happens in the scene, the “weapon focus effect”, was found to be stronger for negative scenes (Loftus et al., 1987). This affect which was also named as the “attentional narrowing hypothesis”, explains the direct attention on the central details due to the emotional intensity of the content which overshadows the peripheral details (Christianson, 1992). These studies bolstered the importance of emotion in cognitive processing apart from the low level and high level factors which were explained in previous chapters.

Studies of emotional content typically rely on people’s ratings of the strength of the emotion in the image (see, for example, the IAPS dataset; Lang, Bradley & Cuthbert, 2008). Recent advances in easy to use arousal measurement techniques, such as heart rate (Appelhans & Luecken, 2006) and skin conductance (Lykken & Venables, 1971) have led researchers to more strongly rely on neurophysiological responses. In the study, to reinforce the assessment of the relationship between the memory performance and the expertise on emotional stimuli,

confidence ratings are collected indicating how confident participants were that they had seen the exact same scene before. The use of confidence ratings in the context of memory was inspired by the work by Kensinger and Schacter (2006), who found that confidence ratings were higher for positively emotional material irrespective of the accuracy of the memory responses. Therefore, I will rely on both types of measures to examine the association between conscious and unconscious responses to emotional stimuli.

In measuring emotional responses, I will also make use of pupil diameter, which has been shown to relate not just to cognitive processing (Beatty, 1982; Beatty & Lucero-Wagoner, 2000 for a review; Mathot et al., 2015), but also to emotional processing (Bradley et al. 2008; Partala, & Surakka, 2003). For example, emotionally arousing pictures (either negative or positive) elicited a larger pupil response than neutral images, which covaried with skin conductance responses (Bradley et al., 2008). Likewise, pupil responses were larger to auditory stimuli that were rated positively or negatively than stimuli that were considered to be neutral (Partala & Surakka, 2003). When interpreting pupil responses, however, it is important to take into account a few confounding factors. For example, immediately after stimulus onset, pupil diameter increases, known as the light reflex. The initial pupil response should therefore be discounted when considering the pupil response due to emotion (Bradley et al., 2008). Moreover, pupil diameter is linked to properties of the upcoming saccade target (Mathot et al., 2015), and therefore pupil responses immediately before saccades need to be interpreted with caution.

While many studies have examined the role of emotional content on attention and memory, few studies have examined how expertise influences these effects. Studies that examined the role of expertise have largely focused on music or the arts. For example, one study compared

the emotional and aesthetic abstract artwork judgements of experts and novices and found that while experts rated the artworks higher on aesthetic aspects, but did not show a similar enhancement of emotional ratings (Van Paasschen, Bacci & Melcher, 2015). Likewise, Castro and Lima (2014) found that experts rated the emotional contents of music more accurately than novices. In contrast, enhanced emotional responses to music in experts was not found by Song, Dixon and Pearce (2016), but this could be due to a difference in the criterion considered (judged emotion versus experienced emotion).

As mentioned before, in eyewitness research, negative emotion has been found to increase the “weapon focus effect” (Loftus et al., 1987). One study extended the study of this effect to police officers, and found that this group of participants were not susceptible to the effect (Hulse & Memon, 2006), suggesting that expertise with crime related situations may protect against the influence of emotion on cognition, although a motivational bias may have played a role.

With the above in mind, the present study aimed to study the role of emotion in the crime scene investigation process. A broad spectrum approach was taken by investigating the role of expertise on a range of measures, including pupil response, eye movements and memory, in order to enhance the chances to find effects of expertise, and to gain a better understanding of how expertise modulates the effect of emotion on cognition, if found. Based on earlier studies, experienced CSIs are predicted to show weaker emotional responses to crime scenes and weaker influences of emotional contents on attention and memory.

7.2. Method

7.2.1. Participants

Twenty-two participants (18 women and 4 men, ages ranged from 19 to 38) took part in the study. Eleven participants were the undergraduate forensic students who had training in CSI (2 male; 9 female; age $M = 24.09$ years, $SD = 6.12$). Eleven participants were the students or researchers from other departments without training in CSI or forensic sciences (2 male; 9 female; age $M = 23.82$ years, $SD = 4.19$). Participants were reimbursed for their time by receiving £5 for their participation.

7.2.2. Ethics Statement

The Ethics committee of the School Psychology of University of Lincoln approved the protocol of the study. Participants all provided written consent and were provided with written instructions and were verbally debriefed at the end of the study.

7.2.3. Materials

A set of negative and neutral images from the Nencki Affective Picture System (NAPS, Marchewka, Zurawski, Jednorog & Grabowska, 2014) and International Affective Picture System (IAPS; Lang et al., 2008) and a set of crime scene images were used as stimuli in the study. Crime scene images were obtained from a range of newspaper websites, such as the Daily Mail. All images were rescaled to a size of 640x480 pixels (landscape) or 480x640 pixels (portrait) resolution, equalling approximately 17 by 13 degrees of visual angle.

Because studies have shown that people in scenes strongly attract the observer's gaze (Birmingham, Bischof & Kingstone, 2008a; Birmingham et al., 2008b; Birmingham et al., 2009) and increase emotion evaluation of the images (Löw, Bradley & Lang, 2013), only images without people or bodies were used. Not using images with bodies also reduced the chances of people being highly disturbed by the crime scene images.

While emotion ratings were available for the NAPS and IAPS images, no such ratings existed for the crime scenes. The 70 crime scene images downloaded from the newspaper websites were therefore pretested for valence and arousal using a web based questionnaire presented to 36 participants (32 male, 4 female, ages ranging from 22 to 80, $M = 34.53$, $SD = 10.35$). Participants were asked to rate the images' valence and arousal dimensions on a 9 point Likert scale. Mean rating scores were ranged from 1.81 ($SD=1.39$) to 5.65 ($SD=2.29$) for valence and 3.81 ($SD=2.45$) to 7.47 ($SD=2.32$) for arousal.

For the main experiment, 30 images were selected from the pretested crime scene images, half of them were evaluated as negative and the other half of them as neutral with varying in arousal ratings. The same numbers of images were taken from the NAPS and IAPS databases. The mean (SD) valence/arousal ratings for the overall negative images were (2.73 (0.33) / 6.21 (0.55)) and for the neutral images were (4.62 (0.32) / 4.73 (0.29)). Paired samples t-tests showed that the ratings in valence ($t(14)=-15.88$, $p < 0.001$, $d=-5.81$) and in arousal ($t(14)=8.39$, $p < 0.001$, $d=3.37$) were significantly different across the two groups of images (similar results were obtained using the non-parametric Wilcoxon Signed Rank Test).

Studies have shown that environmental illumination and the luminance of the stimuli influences pupil size variations (Lang et al., 2008). To incorporate such influences in the

results, the average luminance of each image was calculated using the RGB values (Geangu, Hauf, Bhardwaj, & Bentz, 2011), by applying the equation: $\text{luminance} = 0.2126 \times R + 0.7152 \times G + 0.0722 \times B$. Comparisons of the different image groups on this measure did not show luminance differences for negative ($M=98.64$, $SD=16.54$) compared to neutral ($M=100.72$, $SD=10.59$) images ($t(14)=-0.37$, $p = 0.72$, $d=-0.15$). Detailed information about the ratings and measurements of the images can be found in Appendix C.

7.2.4. Apparatus

Stimuli were presented on a LCD computer 22 inch monitor with a 1680x1050 pixel spatial resolution at a 59 Hz sampling rate. An Eyelink 1000 desk mounted system was used to record eye movements at a sampling rate of 1000 Hz. The centroid mode was used for the pupil area recording over the ellipse mode algorithm as recommended by the manufacturers of Eyelink. In Centroid mode, the number of black pixels on the image of the pupil area were computed and given in a special measurement, “arbitrary units (au)”, as a result. Participants were seated at a viewing distance of 60 cm from the computer monitor with their head in a head and chin rest. The experiment was set up using Experiment Builder and the eye tracking data was extracted using the DataViewer software and analyzed using custom-built MATLAB scripts.

7.2.5. Procedure and Design

Before the experiment, participants were given an introduction to the study and the eye tracking equipment. The experiment started with task instructions (they were going to view two sets of images, crime scenes and natural scenes, for a recognition task) and an initial

practice session, during which the researcher provided assistance to ensure that the participants got familiar with the eye tracker and understood the task requirements.

To ensure correct tracking of individuals' eye movements, the eye tracker was calibrated using the default nine point calibration procedure before the presentation of the stimuli. Calibration was considered successful if the nine randomly illuminating calibration points superimposed with eye position and was reported accurate by the Eyelink software which was approximately associated with an approximate 0.5° accuracy of visual angle on the screen. Otherwise, the calibration process was repeated. After calibration, participants completed an initial practice session which consists of negative and neutral image examples with feedback from the examiner to understand the task in the experiment and to get familiar with the eye tracker and the task process. This practice session also provided the opportunity to ask questions. The practice session involved exactly the same procedure of the real study, containing a total of three images.

In the first session, participants viewed the stimuli while recording their behavioural (pupil dilation) responses. The participants were told not to look away from the computer screen and view the images. Each trial started showing a greyscale background (RGB, 110,110,110) for 2 seconds before image presentation. Afterwards, the image was presented for 6 seconds and this was followed by an inter trial interval between images for 10 seconds while showing another grey background (see, Bradley et al., 2008). Pupil and eye movements were sampled for 2 seconds prior to image onset, for 6 seconds during image onset, and 3 seconds following image offset. After viewing 20 images, a short break was given to participants. No more than two images from the same emotional content were presented consecutively. Images were presented in a random order for each participant.

After all the trials were completed, an immediate recognition test was performed. During the recognition test, 120 images (60 old images and 60 mirror-reversed versions of old images) were presented to the participants for 2 seconds each (c.f., Riggs, McQuiggan, Anderson & Ryan, 2010) in two blocks. Each block had equal numbers of each of the four conditions (negative versus neutral and crime scene versus natural scene images). No more than two images from the same emotional content or image type were presented consecutively. Participants were asked to indicate for each image whether it was the original from the experimental block, or the mirror-reverse version, and to rate their confidence in their answer on a 5 point Likert scale using the computer mouse. No feedback was provided during this memory task, and participants could take as much time as needed to complete the task.

After the memory task and a short break, participants were shown with the original images a second time and were asked to evaluate the images' valence and arousal ratings using the same procedure as used in the pre-test of the images, but now presented on the eye tracker's computer screen. In addition to these two dimensions, participants were also asked to rate the images for complexity, as past studies have suggested that this factor can influence attention (e.g., Chainay, Michael, Vert-pré, Landré, & Plasson, 2012). After completing all tasks, participants were debriefed, thanked and reimbursed for their participation. The session took approximately an hour to complete.

7.2.6. Data Analysis

To analyze the eye movement data for pupil responses, blinks were first identified using the default settings of the Eyelink software. Samples up to 100 ms prior and after the blinks were defined as the blink artefacts and removed from the data (Jainta, Hoorman & Jaschinski,

2008). Linear interpolation was then used to estimate the pupil size during blinks and blink artefacts. Data were subsequently smoothed using a 11 point moving average. Before computing peak and average pupil responses, blinks and blink artefacts were removed from the data (Einhäuser, Stout, Koch & Carter, 2008; Gingras, Marin, Puig-Waldmüller & Fitch, 2015).

As a baseline, the average pupil responses from the periods before and after stimulus presentation were used. Pupil response was defined at the maximum pupil size relative to the baseline value (Nuthmann & Van der Meer, 2005). Other measures analyzed were the minimum pupil size (relative time) and the time until peak pupil size. To examine the initial light reflex, the peak dilation of the pupil in the first two seconds was used as a comparison. Overall averages were obtained by averaging per participant averages of the various analysis. These per participant averages also served as the input for the various statistical analyzes. Most analyzes involved mixed ANOVAs testing the effects of Experience (between subjects) and Emotional Content (Negative or Neutral; within subjects factor) and Image Type (Crime Scene or Database Scene; within subjects factor). Pillai's Trace statistics will be reported.



Figure 7.1. Example of Negative Stimuli. Right: Crime scene image from lab developed stimuli Left: Image from one of the popular affective image databases (NAPS)



Figure 7.2. Example of Neutral Stimuli. Right: Crime scene image from lab developed stimuli Left: Image from one of the popular affective image databases (NAPS)

The data were inspected for violations of normality assumption by the examination of standardized residuals, and by means of a Shapiro-Wilk test. When the normality assumption was violated, square root transformations were applied to the data before computing the relevant statistics. The average data plotted, was not square transformed, as this would make the data plots difficult to interpret.

7.2.7. Results

7.2.7.1. Pupil Responses

Peak pupil responses are shown in Figure 7.3, separately for emotional content, image type, and experience groups. A mixed 2 (emotional content; negative, neutral) by 2 (image type; crime scene, database) by 2 (experience; forensic, non-forensic) factors ANOVA indicated that there was no effect of experience ($F(1,20)=1.22$, $p=0.28$, $\eta^2 = 0.06$) and emotional content ($F(1,20)=0.53$, $p=0.48$, $\eta^2 = 0.03$) on peak pupil response. A main effect of image type was found ($F(1,20)=6.86$, $p=0.016$, $\eta^2 = 0.26$; Larger pupil distribution while viewing

crime scene images ($M=150.88$) compared to database images ($M=136.86$) in general). No interactions were found between experience and emotional content ($F(1,20)=0.33$, $p=0.86$, $\eta^2 = 0.002$), experience and image type ($F(1,20)=0.05$, $p=0.83$, $\eta^2 = 0.002$) and experience and emotional content and image type ($F(1,20)=0.44$, $p=0.52$, $\eta^2 = 0.02$). This suggests that peak pupil responses did not show different distributions in each emotional content of the image types.

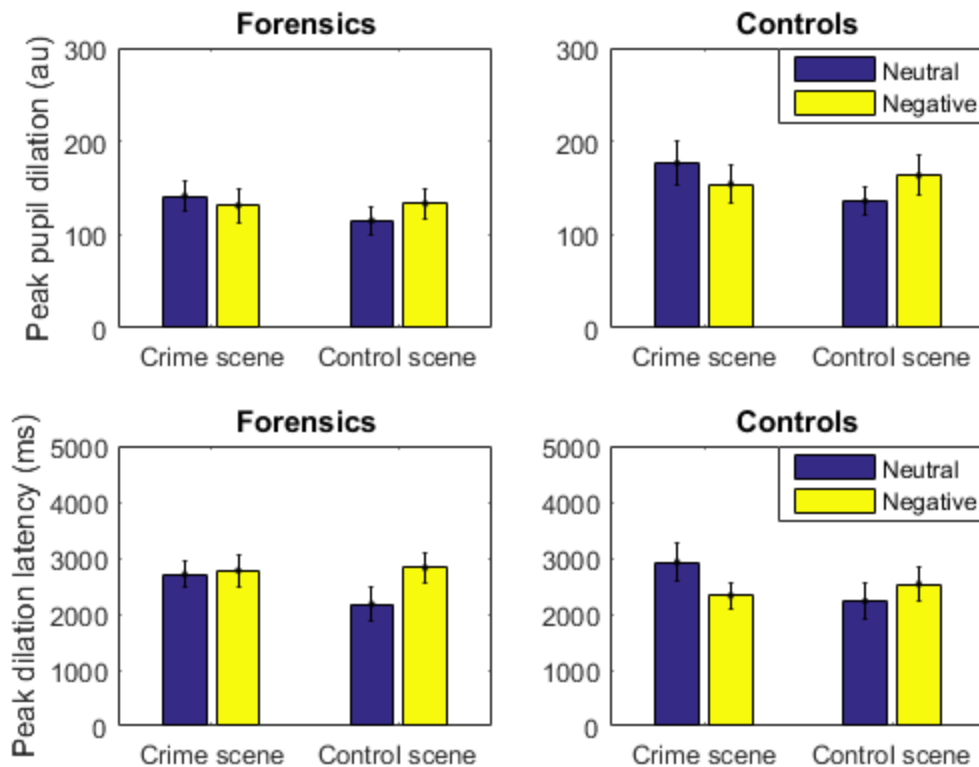


Figure 7.3. At the top: Average Peak Dilations (au), at the bottom: Average Peak Dilation Latency (ms) for each emotional content on image types per participant groups. (\pm SE, $N=22$)

To examine the speed of onset of the pupil response, the latency to peak dilation was investigated. Mean time taken to peak dilation are shown in Figure 7.3, separately for each emotional content on each image type for experience groups. The results showed no effect of experience ($F(1,20)=0.1$, $p=0.76$, $\eta^2 = 0.05$) and emotional content ($F(1,20)=2.02$, $p=0.17$,

$\eta^2 = 0.09$) on time taken to peak dilation of the pupil. A main effect of image type was found ($F(1,20)=4.72$, $p=0.04$, $\eta^2 = 0.19$; Larger times while viewing crime scene images ($M=2.68$ s) compared to database images ($M=2.44$ s) in general). There was an interaction between experience and emotional content ($F(1,20)=11.35$, $p=0.003$, $\eta^2 = 0.36$; Shorter time responses on peak dilation for forensic students on neutral images ($M=2.44$ sec; for inexperienced group $M=2.58$) and higher for negative images ($M=2.8$ s; for inexperienced group $M=2.43$ s) in the absence of experience and image type ($F(1,20)=0.0$, $p=0.99$, $\eta^2 = 0.00$) and experience and emotional content and image type ($F(1,20)=0.67$, $p=0.42$, $\eta^2 = 0.03$) suggested that time to peak dilation in pupil responses between groups did merely show different distributions on the emotional content of the stimuli.

To examine the influence of the pupil light response on the data, the first 2 seconds of the response were analyzed separately. For this time period, peak dilation showed no effect of experience ($F(1,20)=0.44$, $p=0.52$, $\eta^2 = 0.02$). Nor was there an effect of emotional content ($F(1,20)=1.33$, $p=0.26$, $\eta^2 = 0.06$) or image type ($F(1,20)=1.28$, $p=0.27$, $\eta^2 = 0.06$). No interaction was found between experience and image type ($F(1,20)=0.03$, $p=0.87$, $\eta^2 = 0.001$), experience and emotional content ($F(1,20)=2.71$, $p=0.12$, $\eta^2 = 0.12$) and between experience, emotional content and image type ($F(1,20)=1.04$, $p=0.32$, $\eta^2 = 0.05$). This suggests that time to peak dilation in pupil responses between groups were not affected by the emotional content or image type. Mean peak dilation are shown in Figure 7.4, separately for each emotional content on each image type for experience groups.

Analysis of the latency to peak dilation in the first 2 seconds showed no effect of experience ($F(1,20)=0.13$, $p=0.73$, $\eta^2 = 0.006$). However, there was a significant effect of emotional content ($F(1,20)=5.11$, $p=0.035$, $\eta^2 = 0.2$; Shorter time to peak dilation for emotional

($M=0.69$) compared to neutral pictures ($M=0.76$) in general) and image type ($F(1,20)=20.69$, $p<0.001$, $\eta^2 = 0.51$); Longer time to peak dilation for crime scene ($M=0.81$) compared to database images ($M=0.77$) in general). Also there was an interaction between image type and emotional content ($F(1,20)=13.37$, $p=0.002$, $\eta^2 = 0.4$). No interaction was found on experience and image type ($F(1,20)=0.82$, $p=0.38$, $\eta^2 = 0.04$), nor was there an interaction between experience and emotional content ($F(1,20)=3.54$, $p=0.23$, $\eta^2 = 0.07$) or between experience, emotional content and image type ($F(1,20)=0.06$, $p=0.81$, $\eta^2 = 0.003$). This suggests that time to peak dilation in pupil responses between groups show different distributions on the emotional content of the each of the image type. Mean peak dilations are shown in Figure 7.4.

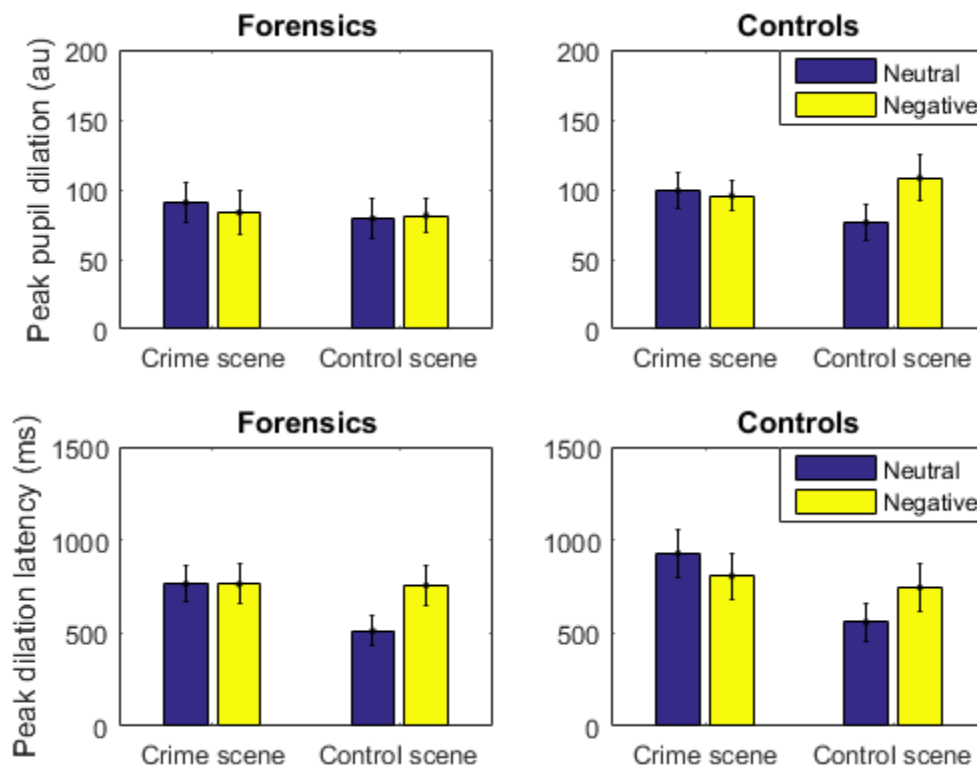


Figure 7.4. At the top: Average Peak Dilations (au). At the bottom: Average Peak Dilation Latency (ms) for the initial 2 seconds (after image onset) on image types for each emotional content per participant groups. (\pm SE, $N=22$)

7.2.7.2. Eye Movement Characteristics

The data analysis was performed using the eye movement measures from Dietz, Bradley, Okunc & Bowersa (2011). Mean number of fixations are shown in Figure 7.5, where for each image and content on each participant group the average number of fixations are shown. The results revealed that emotional content slightly affected the number of fixations ($F(1,20)=3.56$, $p=0.074$, $\eta^2 = 0.15$) and the effect was marginally significant (More fixations were made when viewing emotional ($M=17.65$) compared to neutral ($M=17.29$) pictures). No differences were found on the effect of experience ($F(1,20)=0.7$, $p=0.41$, $\eta^2 = 0.03$) or image type ($F(1,20)=1.3$, $p=0.27$, $\eta^2 = 0.06$), nor were there any interaction between experience and image type ($F(1,20)=1.05$, $p=0.32$, $\eta^2 = 0.05$), experience and emotional content ($F(1,20)=0.19$, $p=0.67$, $\eta^2 = 0.01$) or among experience and image type and emotional content ($F(1,20)=1.7$, $p=0.21$, $\eta^2 = 0.08$) and it suggested that differences between emotional content and image types in terms of the number of fixations were similar between the participant groups.

Mean saccadic amplitudes are shown in Figure 7.5. Saccadic amplitude was not influenced by experience ($F(1,20)=1.58$, $p=0.22$, $\eta^2 = 0.07$). There was an effect of image type ($F(1,20)=19.31$, $p<0.001$, $\eta^2 = 0.49$; Larger saccadic distribution while viewing crime scenes ($M=4.31$) compared to database images ($M=4.15$) in general). Saccade amplitudes were unaffected by emotional content ($F(1,20)=1.41$, $p=0.25$, $\eta^2 = 0.07$). No interactions were found between experience and image type ($F(1,20)=3.12$, $p=0.09$, $\eta^2 = 0.14$), or experience and emotional content ($F(1,20)=1.5$, $p=0.24$, $\eta^2 = 0.07$). There was an interaction between experience, image type and emotional content ($F(1,20)=5.27$, $p=0.033$, $\eta^2 = 0.21$) indicating saccadic amplitude differences between the participant groups were affected by the image

and the content used. This three-way interaction was due to relatively large saccades amplitudes in inexperienced students for both types of images. Novices also showed larger saccade amplitudes for crime scene, compared to natural images.

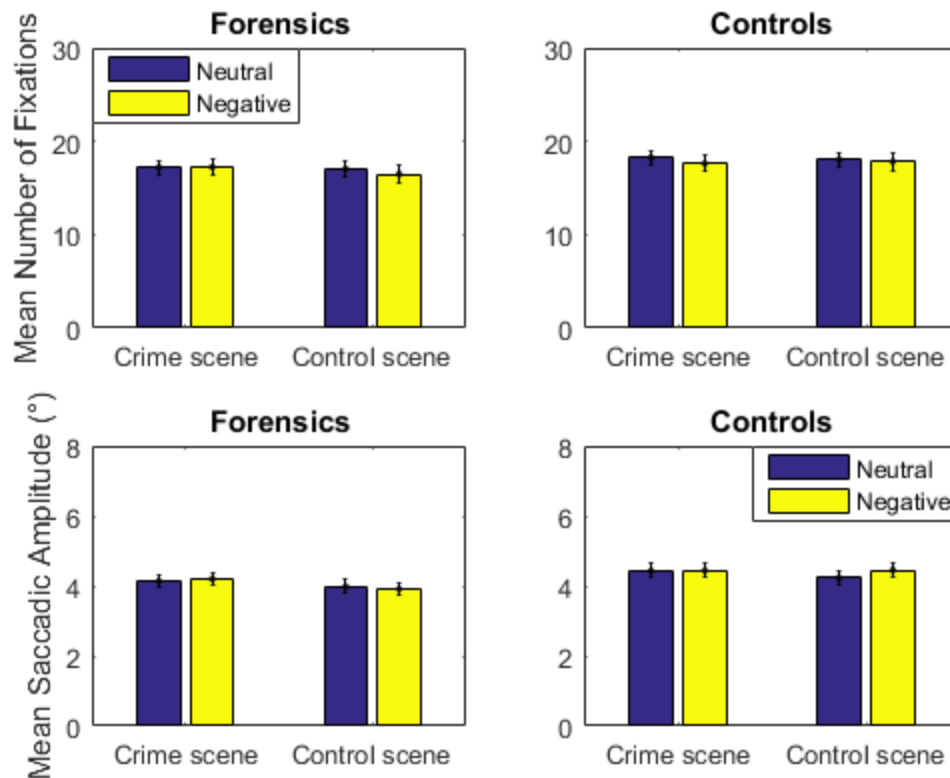


Figure 7.5. At the top: Mean number of fixations, At the bottom: Mean saccadic amplitude (in degrees of visual angle, °) separately shown for the two emotional contents (neutral versus negative), types of scenes (crime scene versus control scene) and participant groups (novices versus experienced CSI students). (\pm SE, N=22)

7.2.7.3. Memory performance

Memory performance was tested by mixed original images with mirror-reversed version of the image and by asking participants to indicate which images were original and which were mirror-reversed. Two blocks were used, but due to technical problems, data from the second

block were not recorded. Fortunately, images were randomized, and therefore there were sufficient trials left for each condition to analyze on the basis of the first block alone. Because proportions correct violated the normal distribution assumption for mixed factor ANOVAs, arcsin transforms were applied to the data before statistical analysis.

Figure 7.6 plots the memory accuracy for the different groups and conditions. No effect of experience was found ($F(1,20)=0.6$, $p=0.81$, $\eta^2 = 0.003$), but there were effects of image type ($F(1,20)=9.1$, $p=0.007$, $\eta^2 = 0.31$) and mirror-reversing on memory performance ($F(1,20)=51.4$, $p<0.001$, $\eta^2 = 0.72$). No interactions were found between experience and image type ($F(1,20)=0.12$, $p=0.91$, $\eta^2 = 0.001$) or experience and manipulation of images ($F(1,20)=0.22$, $p=0.64$, $\eta^2 = 0.01$). However, there was a significant interaction among experience and emotional content and manipulation ($F(1,20)=6.4$, $p=0.02$, $\eta^2 = 0.24$), meaning that memory performance differed between groups within sets of images. It should be noted that these results need to be interpreted with care, as response accuracy may not have met all test assumptions even after transformation.

Figure 7.6 shows the confidence ratings for the memory responses for the different groups and conditions. Confidence about the memory test responses were not affected by experience ($F(1,20)=0.78$, $p=0.78$, $\eta^2 = 0.004$). A main effect of image type ($F(1,20)=8.26$, $p=0.009$, $\eta^2 = 0.29$) and manipulation condition was found ($F(1,20)=17.81$, $p<0.001$, $\eta^2 = 0.47$), in the absence of any interactions between experience and image type ($F(1,20)=0.04$, $p=0.85$, $\eta^2 = 0.002$), or experience and manipulation ($F(1,20)=0.29$, $p=0.59$, $\eta^2 = 0.02$), or between experience, emotional content and manipulation ($F(1,20)=0.39$, $p=0.54$, $\eta^2 = 0.02$).

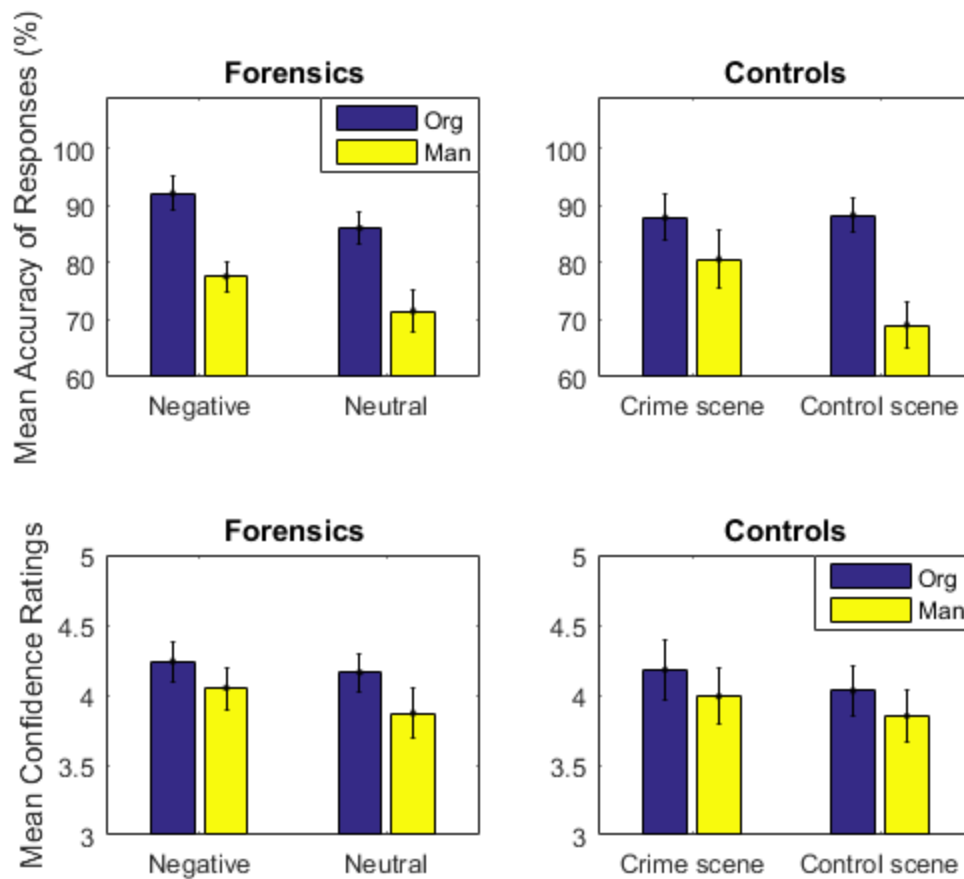


Figure 7.6. At the top: Mean memory accuracy (%), At the bottom: Mean Confidence Ratings separately shown for the two emotional contents (neutral versus negative), types of scenes (crime scene versus control scene) and participant groups (novices versus experienced CSI students). (\pm SE, N=22; in legend, Org=Original Image, Man=Manipulated Image)

7.2.7.4. The Valence and Arousal Ratings

Valence ratings and emotional content were rated separately by external raters (IAPS and NAPS images) and crime scenes (36 independent raters), but also by the participants in the study. Before statistically analyzing these data, the ratings were square root transformed to avoid violations with the normal distribution assumption of the tests applied.

Figure 7.7 plots the valence ratings by the participants in the study. The results indicated no effect of experience on valence ratings ($F(1,20)=2.73$, $p=0.11$, $\eta^2 = 0.12$). There was a main effect of image type ($F(1,20)=74.16$, $p<0.001$, $\eta^2 = 0.79$) and emotional content ($F(1,20)=120.56$, $p<0.001$, $\eta^2 = 0.86$). No interaction was found between Experience and Image Type ($F(1,20)=0.91$, $p=0.35$, $\eta^2 = 0.04$), Experience and emotional Content ($F(1,20)=0.35$, $p=0.56$, $\eta^2 = 0.02$) or Experience, Image Type and Emotional Content ($F(1,20)=0.22$, $p=0.64$, $\eta^2 = 0.01$). This suggests that differences between emotional content and image types in terms of the valence ratings were similar across the participant groups.

Figure 7.7 shows the arousal ratings for the different groups and conditions. Arousal ratings did not show an effect of experience ($F(1,20)=0.31$, $p=0.58$, $\eta^2 = 0.15$). A main effect of image type was found ($F(1,20)=30.85$, $p<0.001$, $\eta^2 = 0.61$), in the absence of an effect of and emotional content ($F(1,20)=0.04$, $p=0.85$, $\eta^2 = 0.002$) and experience. There was an interaction between experience and image type ($F(1,20)=6.43$, $p=0.02$, $\eta^2 = 0.24$), but no interactions were found between experience and emotional content ($F(1,20)=1.19$, $p=0.29$, $\eta^2 = 0.06$), or among experience, image type and emotional content ($F(1,20)=0.19$, $p=0.67$, $\eta^2 = 0.01$). The experienced group (forensic students) evaluated the images as less arousing than the inexperienced group.

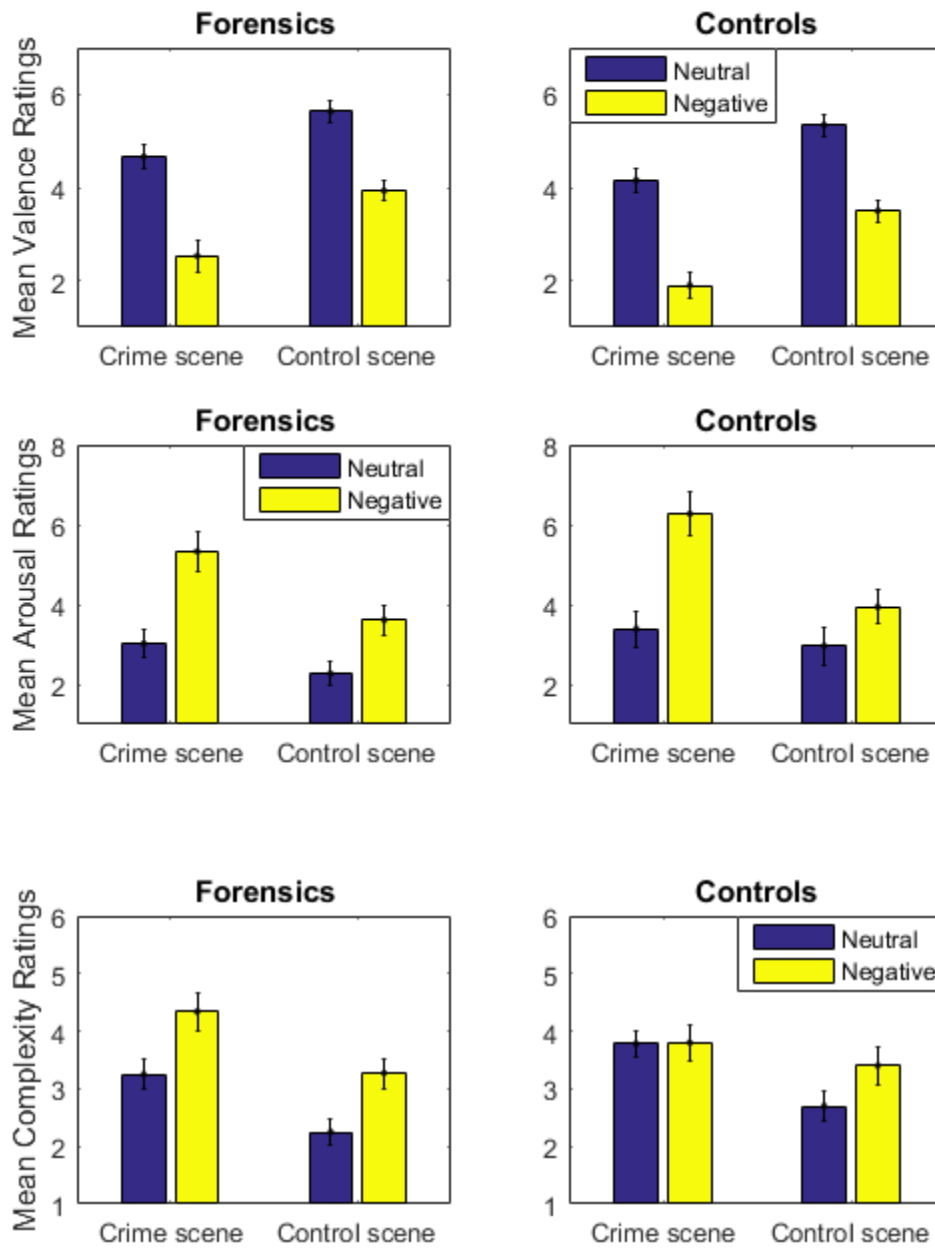


Figure 7.7. At the top: Mean Valence Ratings, Middle: Mean Arousal Ratings, At the bottom: Mean Complexity Ratings, shown separately for the two emotional contents (neutral versus negative), types of scenes (crime scene versus control scene) and participant groups (novices versus experienced CSI students). (\pm SE, N=22)

Figure 7.7 plots the rated complexity across groups and conditions (past research has suggested that complexity may be a factor in image memory; Chainay et al., 2012). No effect

of experience on complexity ratings were found ($F(1,20)=0.17$, $p=0.68$, $\eta^2 = 0.01$). There were main effects of image type ($F(1,20)=70.16$, $p<0.001$, $\eta^2 = 0.78$) and emotional content ($F(1,20)=24.11$, $p<0.001$, $\eta^2 = 0.55$). No interaction was found between experience and image type ($F(1,20)=1.81$, $p=0.19$, $\eta^2 = 0.08$). Also there was no interaction between image type and emotional content ($F(1,20)=2.24$, $p=0.15$, $\eta^2 = 0.1$). There was a significant interaction between experience and emotional content ($F(1,20)=5.72$, $p=0.027$, $\eta^2 = 0.22$) and a marginal interaction across experience and image type and emotional content ($F(1,20)=3.48$, $p=0.077$, $\eta^2 = 0.15$).

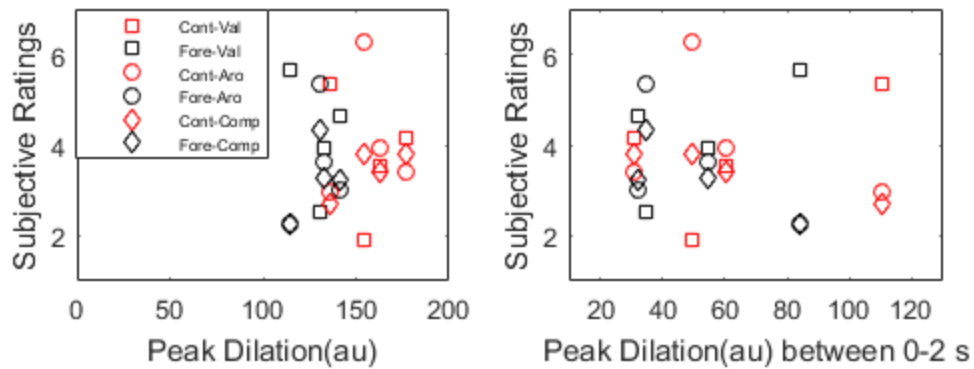


Figure 7.8. Scatter plot comparing peak dilations and subjective ratings for a range of ratings overall (valence (Val), arousal (Aro) and complexity (Comp)). Data are shown for the entire image presentation interval and the first two sections. Black and red symbolic indicate Forensic Students (Fore) and Inexperienced (Cont).

7.3. Discussion

Crime scene investigation is known to induce stress in CSIs. How such stress influences attention and memory, and how experience with crime scene investigation affects these influences is unknown. The present study was the first trying to answer these questions.

Crime scenes and natural images were presented, while eye movements, pupil responses, memory performance, and image ratings were collected. Images were selected to either represent negative emotional scenes or neutral scenes (as rating by independent raters).

Peak dilation and time to peak were unaffected by experience and emotional content. These results are at odds with past studies reporting mean pupil size (e.g. Bradley et al., 2008; Dietz et al., 2011; Pattala & Surakka, 2003). The reason for using peak dilation in the present study was to reduce influences of cognitive load, which is known to more strongly influence average dilation than peak dilation. (Nuthmann & van der Meer, 2005), but is unclear whether this may have influenced the results. Other factors may have played a role. These include the selection of the images and the study design (Lanata et al., 2012) and the initial light reflex (Bradley et al., 2008; Lanata et al., 2012; Schienle, Übel, Gremsl, Schöngassner & Körner, 2016). The analysis of the pupil response in the first two seconds after image onset showed results consistent with the literature (Nuthmann & van der Meer, 2005; Bradley et al., 2008), although small variations in the images presented might have led to the observed differences (Schienle et al., 2016). Other studies have applied grey scale images to avoid influences of the light reflex (e.g., Bradley et al., 2008; Dietz et al., 2011). In the present study such an approach was not feasible as grey scale images of blood stains are unlikely to have the same emotional contents as colour images. Moreover, previous studies have examined pupil responses with colour images and found similar pupil responses when compared with grey scale images (Schienle et al., 2016). In addition, it might be considered that biases with respect to emotional content may change the course of a trial, where positive and harm aspects of a scene are equally likely fixated early in the trial (first 500ms), while harm stimuli are avoided later in the trial (Calvo & Avero, 2005). Such early approach and late avoidance behaviours may have occurred in the present study, although more data would be required to

reliably investigate such effects on pupil diameter measurements across such small intervals (500ms). Furthermore, recent work by Mathot et al. (2015) has found that in goal driven tasks, arousal was not a factor for pupil dilation. Instead pupil dilation was more strongly influenced by cognitive workload. This means that the null findings in the present study can both be due to the higher cognitive workload or due to task design and instructions.

Another reason why no effects of experience on pupil diameter were found, could be the age range of the participants in the two groups. Different studies in emotional information processing showed that age has an important influence on pupil diameter (e.g., Isaacowitz, Wadlinger, Goren, & Wilson, 2006a; Isaacowitz et al., 2006b; Isaacowitz, Toner, Goren, & Wilson, 2008). These studies have suggested that older participants respond more strongly to positive emotions than younger participants. Variations across participants were also found to be larger between younger than between older participants (e.g., Allard, Wadlinger, & Isaacowitz, 2010). Other contributing factors are age (e.g. Gavazzeni et al., 2008), gender differences (e.g. Bradley, Codispoti, Sabatinelli, & Lang, 2001; Glaser, Mendrek, Germain, Lakis & Lavoie, 2012), as these factors influence how strongly people respond to emotional contents. These factors, however, are not likely to have had an effect. Participant groups were balanced on gender and age, and rated images similarly on arousal and valence.

Saccade amplitudes were also very similar across participant groups, and therefore did not show an effect of experience. Novices showed slightly broader scanning patterns (larger saccade amplitudes) than more experienced participants (forensic students) for negative images, but the type of image did not influence this pattern. The smaller saccade amplitudes for the experienced groups could be an indicator of a particular search pattern, however in the literature experts tend to have broader scanning patterns (e.g., Krupinski et al., 2006;

Reingold et al., 2001), so it is unclear what the present results mean. The memory task may have had an influence as well. Experienced participants may have known better how to perform this task, and may have adopted different scanning techniques than the novice participants.

Number of fixations, consistent with the findings in literature (e.g. Bradley et al., 2011; Carniglia, et al., 2012), was influenced by emotional content. To process the emotionally negative stimuli visually, participants from both groups (experienced and inexperienced) applied relatively more distributed sampling patterns.

No effect of experience was found on the valence and arousal ratings. There was an interaction between the experience level and the image type on arousal perception, suggesting that the forensic students rated images lower on arousal overall. This might be an indicator of familiarity with the crime related material. However, valence ratings did not differ across groups. It is unclear why. Related work in arts and music also showed inconsistent effects of experience. For example, while Van Paasschen et al. (2015) found out similar emotion ratings in experts and novices, Castro and Lima (2014) did find influences of experience on emotion perception. These results however, contrasted with Song et al. (2016) who did not find any effect of experience on emotion perception.

No effects of experience on memory were found either. These results are at odds with the earlier findings as expertise effects were found in radiology (Evans et al., 2011; Evans et al., 2016), but memory retention intervals in these latter studies were short, and only memory effects were found for domain related images. As discussed in the previous chapters, images of visual input in sports and radiology is more specific than for crime scene investigation,

and the broad range of visual input in crime scenes may have influenced the results. Crime scenes mostly contain objects one would encounter in day-to-day life, and one may therefore argue that crime scenes are visually more similar to the control images in Evans et al. (2011) and Evans et al. (2016), for which no expertise effects were found.

Despite the large variability in the visual layout of the images and the difficulty of the task (detecting mirror-reversions), memory performance was relatively high (between 70% and 90% with chance level at 50%). Studies on change blindness, in which participants have to detect a change in two subsequently presented images, have shown that performance is linked to the saliency of the change and the amount of clutter in the scene. Because images from the present were rather cluttered and the saliency of the change was small, performance would have been expected to be low, and it is unclear why performance was higher than expected. The high overall performance, however, would have given sufficient space for expertise effects (i.e., they would not be limited by floor effects) and therefore the data suggest that expertise truly has no effect on memory performance.

A possible reason for the high memory performance and the lack of differences between the groups may have been that participants were made aware at the start of the experiment that their memory for the images was going to be tested. Improved memory performance under such circumstances have been consistently reported in the literature, and the effect became known as the effect of intent (Block, 2009). This effect has been attributed to more effortful processing of the images, compared to when no advance warning of the memory test is used (Chainay et al., 2012).

One surprising result in the memory data was the lower performance (albeit still above chance level) of both groups on the manipulated images than the original images. This finding, however, is in line with results by Lawson (2004), who also found worse performance on reversed images in an intentional memory task. Image based approaches assume that any changes in the previously viewed images or scenes (e.g., size, rotation, occlusion) do not impair performance in the memory task (Biederman, 2000; Lu & Liu, 2008; Lu & Liu, 2009; Poggio & Edelman, 1990). Some past work has shown influences of mirror-reversion on self-face recognition (e.g., Smith, Grabowecky & Suzuki; 2008), but these effects could be very specific to the recognition of an image of one's self, as people normally only see the mirror-reversed image of themselves when looking in the mirror. Exposure to the original layout only occurs in photographs, which may be encountered less often than mirrors.

A further finding in the memory data is that responses to negative images were more accurate than responses to neutral images. These improved negative image responses were independent of expertise and the type of scene, and therefore appear to relate uniquely to the images being negative. Improved memory for negative images, compared to neutral images, has been found in a range of studies (Kensinger, Garoff-Eaton, & Schacter, 2006; Humphreys et al., 2010).

The study had several strengths. First, a large number of stimuli were used, either from databases that had been extensively tested and evaluated (NAPS, IAPS), or from a set of crime scenes that were pre-tested for emotional content. Second, the study applied a broad range of measurements to evaluate the effect of emotional content, including, emotion and confidence ratings, pupil dilations, eye movement parameters, and memory performance. If

expertise had a clear effect on cognitive responses to emotional contents, this would have been evident in at least one of the measures. Third, participant groups were closely matched in their gender and their age, meaning that if any differences between groups would have arisen, they would have not been due to these participant characteristics.

The use of pupil responses was a particular strength of the study. Studies have shown that pupil size is extremely difficult to control consciously, despite the many websites claiming the opposite. Any expertise effects on pupil responses should therefore be unrelated to intent of the participants. Pupil responses, however, have also been found to vary substantially across participants (Beatty & Lucero-Wagoner, 2000), which may have limited the statistical power to detect group difference. In part, these individual differences can be explained by personality traits (Schienle et al., 2016), which were not controlled for in the present study. Future studies, however, should use scales to measure personality traits, to account for any variance in pupil responses that could possibly arise from these measures.

Despite these strengths, there were also shortcomings. The sample size was relatively low with 11 participants in each group. Participants were financially reimbursed for their participation, and still it was difficult to achieve high participant numbers. Part of the recruitment issue was due to a limited number of forensic students at the testing site, but also recruitment of control participants proved difficult. A possible reason was that the study was conducted in the period near and during the exams and possible participants may have decided not to take part, because they were too busy preparing for the exams.

Another limitation is that participants were all recruited only from the university, which may limit the extent to which the results can be generalized to the population. The restriction of the

sample to university students and staff, however, had the advantage that groups were relatively homogenous, and should have increased the chances of finding group difference.

A possibly more important limitation is that the forensic participants were all in their second or third year of their forensic studies and had no experience in casework. The experts in the study may therefore not have had the relevant exposure to distinguish themselves from the control groups on the measures used. This contrasts with the studies reported earlier, where actual CSIs formed the expert groups, and it would therefore be beneficial to repeat the study with actual CSI as the experts. This would, however, introduce confounding factors that need to be controlled for, such as the age and gender of the participants, and the testing environment (recruitment is likely to be only possible by moving the setup near to the police headquarters).

7.4. Conclusion

No effects of crime scene expertise were found on the emotional responses to neutral and negative images of crime scenes and natural images, despite the broad range of measures considered (emotion ratings, confidence ratings, eye movement measures, pupil dilation, memory performance). The results suggest that expertise in crime scenes does not influence how emotional contents modulates cognitive aspects (attention and memory) of the crime scene investigation process. Future studies should consider a broader range (CSIs, forensic students, age-matched controls) and larger groups of participants, and should consider how the effects of emotional context affect beyond memory and attention.

Chapter 8. Conclusions

This final chapter will provide an overview of the findings of the studies and will discuss the outcomes in the context of possible applications in crime scene examinations. The chapter will conclude with recommendations for future studies.

8.1. Background

The present work applied eye tracking in crime scene investigation to examine the roles of perception and cognition to the crime scene investigation process. The approach taken was to compare experts and novices in order to determine how expertise in crime scene investigation influences how scenes are perceived and processed. As an application, eye movements may be used in the future to assess the skill of a crime scene investigator (CSI), or to improve teaching of the crime scene investigation process. Despite the neglected importance of human factors in forensic sciences, perceptual and cognitive processes are one of the essential factors for effective and efficient performance of CSIs and the crime scene examination process (Dror, 2015; Julian et al., 2011; Kely et al., 2011). These studies also indicate that human factors need to be considered to improve the quality of the crime scene investigation process, beyond improvements made on the basic technological advances (Kloosterman et al., 2015).

With the changing atmosphere of the investigations in policing (Sheridan, 2013), forensic science has become much more crucial and influential over the last few decades. Factors in this increasing influence have been the introduction and advancement of DNA sequencing,

and the depiction of the crime scene investigation process in popular TV shows. A literature review, however showed that while technical aspects of the crime scene investigation process were often described, only few texts mentioned or described cognitive aspects of crime scene investigation. Such cognitive aspects have started to receive more attention after the NAS Report in 2009, which highlighted the lack of collaboration with cognitive psychologists, compared to other domains. Still, studies on cognitive aspects are far from reaching adequate numbers (Found, 2015). Existing studies have used different research methods to investigate the underlying cognitive processes and the factors that affect different levels of CSI processes ranging from case reports from police and courts (e.g., Julian et al., 2012), observational studies (e.g., de Gruijter et al., 2016) to interviews (e.g., Kelty et al., 2011). For example, in their study, Kelty et al. (2011) provided findings on the required skills of competent CSIs while collecting data from experienced CSIs and filled in an important gap in the domain. The developments in technology also helped to improve these methods as well as combining some of these in the studies although only recently have a combination of these methods been applied in practical environments. As one of few previous studies, Baber and Butler (2012) made use of simulated crime scenes and a head mounted camera with concurrent verbal reporting in their study and traced an actual crime scene search process. In another recent study, de Gruijter, et al. (2016) used a combination of observations and verbal reports in a simulated scene investigation study. Applying combination of similar methods which have been used in different domains (e.g., Jarodzka et al., 2010 in biology) could lead to better crime scene examination performance and more accurate and fair justice outcomes.

Just and Carpenter's eye mind assumption (1980) led researchers to apply eye tracking measures to understand perceptual and cognitive processes. With a better understanding of human eye movements, recent developments in technology, and demonstrations in other

domains, it can be expected that eye movements can provide new insights into perceptual and cognitive processes of professionals in forensic science and also provide new tools to enhance efficiency and effectiveness in this area. Thus, extended research while applying eye tracking methodology can provide valuable findings as an objective tool to measure skills and to probe into underlying mental processes. While scrutinizing human factors, an effective way of learning critical factors (e.g. required skills, attitudes, strategies) is to learn about and from the experts (Ericsson, 2006). Experts' ability to do things easily is different from novices due to learned domain knowledge and extended experience and this is reflected in the differences between experts and novices (Ericsson, 2006).

8.2. Summary of findings.

In this thesis I made comparisons between experts and novices in crime scene examination to understand the underlying mental processes and factors which influence them (Chi, 2006). Across four different studies, this thesis scrutinized the perceptual and cognitive processes in crime scene investigation. Based on previous studies which showed the effect of expertise in other domains using the eye tracking technology (see Gegenfurtner et al. 2011 for a review; also Hermens et al., 2013 for a review on surgery), I expected differences in crime scene processing across levels of expertise. Traditional eye tracking metrics, such as fixation durations and saccades were applied in addition to the analysis of spatial aspects and temporal order effects of saccades. Eye tracking was performed using desk mounted and mobile eye trackers. I also used pupillometry to examine the emotional information processing to gain better understanding of the influence of expertise. Consistent with the empirical findings from other domains on the effect of expertise on eye movement patterns, the assumption was

that there would be top-down influence in eye movements of the experts, while bottom-up (stimulus driven) factors would be the same in expert and novice CSIs.

In the first experiment, I investigated the visual processing and cognitive performance of participants with various levels of training and experience (CSIs, third year forensic students, first year forensic students, and controls aged matched to the CSIs) on the basis of photographs of scenes while using desk mounted eye tracker. The process was very similar to the actual examination process, with respect to assessment and identification of the artefacts which might yield evidential value for further examinations in the forensic laboratories. The use of photographs as stimuli allowed for high resolution and precision eye tracking. While the experiments showed expertise effects on saccadic amplitudes and dwell times on evidence, no expertise effects on visual performance measures was found. Although the eye movement findings were inconsistent with the literature from other domains (see Gegenfurtner et al., 2011 for a discussion), some studies which mainly used static images or video recordings apart from the domain environment demonstrated no difference between the experts and novices (e.g., Sodergren et al., 2010 in surgery). Therefore, the study design (relying on static images with lack of active inspection in crime scene) might have influenced the results in the study. The data also suggested that bottom-up, visual aspects of the scene had larger effects in searching process of the scene, as large effects of the image on both eye movements and performance were found. In contrast to performance data, the verbal report data showed the importance of domain specific knowledge and experience in the domain in line with Schraagen and Leijenhorst (2001). Differences in the verbal reports revealed that experience led to use of different cognitive reasoning strategies (see also, Baber & Butler, 2012). CSIs generally generated more exploratory statements than novices. Experts used their knowledge and experience to consider possible scenarios and adapt the appropriate

mental models. These results suggest that experts and novices may be fairly similar in their selection of evidence (involving low-level visual attention factors), but differ in their interpretation of this evidence (involving the processing and retention of the evidence).

In the second experiment, expertise in CSIs was scrutinized in a "change blindness" study, also applying eye tracking methodology to examine visual processing differences between CSIs and novices. The results showed no differences in change detection performance between CSIs and novice participants, in contrast to past studies examining expertise (Werner and Thies, 2000; Beck et al., 2013). One of the reasons for that might have been the changing characteristics of each of the investigations and cases than those domains which were much more constrained to be assumed about what could be seen. Also, visual skills could have had much more crucial importance in the sports and the radiology than crime scene investigation. Performance was well below 100% and this might be an indicator for lower performances in real case processes. Future studies will be needed for the generalization. However, the findings of the present study were in line with a recent study of Smart et al. (2014), who was unable to find differences between the expert police officers and novices in eye witness identification. The present study, however, did reveal an effect of expertise on fixation durations, which were longer in experts compared to novices, similar to the results in Experiment 1. As the same experts and novices participated across the two experiments this difference in fixation durations could be due to the particular participants in the study, and may not reveal an expertise effect.

The third experiment moved beyond the use of static images, and instead relied on active exploration of a simulated crime scene. The experiment made use of a mobile eye tracker and eye movements were recorded while participants actively engaged with their scene.

Because the setup could not be moved easily from the headquarters to the University of Lincoln Eye Tracking Lab (which would have involved physically moving a simulated crime scene), two groups of crime scene investigators (senior; more experienced, and volume; less experienced) were compared, rather than experts and novices. Given that significant differences were found between volume and senior investigators (fixations on exit and entry points, and objects that were out of place), larger, even more robust differences may be expected when comparing senior investigators and novices. Future studies should examine whether this is indeed the case.

A fourth study examined the influence of expertise on emotional processing of crime scenes. Pictures of crime scenes that were rated highly on emotional content and neutral images were presented to forensic and non-forensic students and pupil diameter, eye movements, memory and emotion ratings were collected. No expertise effects on either of the measures was found. Fairly small sample sizes were used (participants were difficult to recruit and often did not present when recruited), and the difference in expertise between the two participant groups was relatively small (expert CSIs were no longer available to take part). The study also found that pupil diameter did show the expected modulation of stimulus presentation, meaning that the measure is a viable measure for future studies that could look into whether emotional content can be shown to have an effect if larger samples and more diverse expertise groups are used.

8.3. Implications of the Findings

In crime scene investigations, the main task is to reconstruct possible prior activities that explain what happened in the scene and who was the perpetrator (Jackson et al., 2015). This

dissertation focused on the first stage of the crime scene investigation process: The inspection of the scene and collection of evidence. The findings obtained fit in the information-processing approach described by Wickens and Holland (2000), demonstrating how clues in the scene are combined with information in memory to construct the most likely scenario of what happened (see Figure 8.1). The small difference between experts and novices and the less than perfect performance on change blindness and memory tasks, however, also highlight the capacity limits of cognition. It is important to acknowledge these limitations on the cognitive system, as they can strongly influence the investigation process (Dror, 2015; Lee & Pagliaro, 2013; Schiffer, 2009)

It is generally accepted that not only bottom-up (stimulus-driven) factors influence the crime scene investigation process. Top-down factors, such as expertise and expectations play an additional important role (Dror, 2011). The present results show large effects of bottom-up factors (image, scenario) on eye movements. One of the top-down factors (expertise) had limited effects, but that does not mean that top-down factors do not influence eye movements. For example, expectations may drive experts and novices alike and drive where people look when approaching a crime scene (applying expectations about crime scenes; Lee & Pagliaro, 2013). While eye movements provided little evidence of differences between experts and novices, the verbal reports suggested that experts process evidence differently (in particular when CSIs were compared to students). Specifically, experts appeared to be better at linking the evidence to the case and at incorporating domain knowledge and previous experience (Schiffer, 2009). A similar pattern of results is found in medical domains, with top-down effects mostly on reasoning and decision making processes (Corcoran, 1986; Norman, Coblenz, Brooks, & Babcock, 1992), although some stronger effects on eye movements have been found (e.g., in surgery, Hermens et al., 2013).

Although the present work has not found strong expertise effects on eye movements, the technique provides unique insight into how crime scene investigators and other experts (e.g., surgeons) approach a case. Eye movements have been studied in the context of domain expertise in surgery (Law et al., 2004; Kocak et al., 2005; Tien, Atkins, Zheng, & Swindells, 2010; Wilson et al., 2011, Wilson, Coleman, & McGrath, 2010; Vine, Chaytor, McGrath, Masters, & Wilson, 2013), radiology (Nodine & Kundel, 1987; Kundel & La Follette Jr, 1972), aviation (Fitts et al., 2005), chess (Charness et al., 2001; Reingold et al., 2001) and various sports (Shank & Haywood, 1987; Moreno, Reina, Luis, & Sabido, 2002; Takaakikato & Fukuda, 2002). In forensic sciences the application of eye movements has been limited, but this dissertation shows that there is potential in their use (see also, Baber and Butler, 2012). Although I did not directly compare the retrospective think aloud protocol with and without eye movements, conversation with the participants suggested that they found the addition of their eye movements to the play-back of the scene useful. Such an interpretation would be in line with previous work investigating the use of eye movements in think aloud protocols (Cooke & Cuddihy, 2005; Elling et al., 2011; Gerjets et al., 2011; Hyrskykari et al., 2008)

For example, recordings of eye movements were beneficial in the cued retrospective report, providing a unique stimulus for recalling one's lines of thought when approaching a crime scene. Eye movements, used in this way, can complement qualitative methods (e.g., Helsloot & Groenendaal, 2011; de Gruijter et al., 2016). Future uses of eye tracking can be to assess cognitive workload and how people deal with various levels of task complexity.

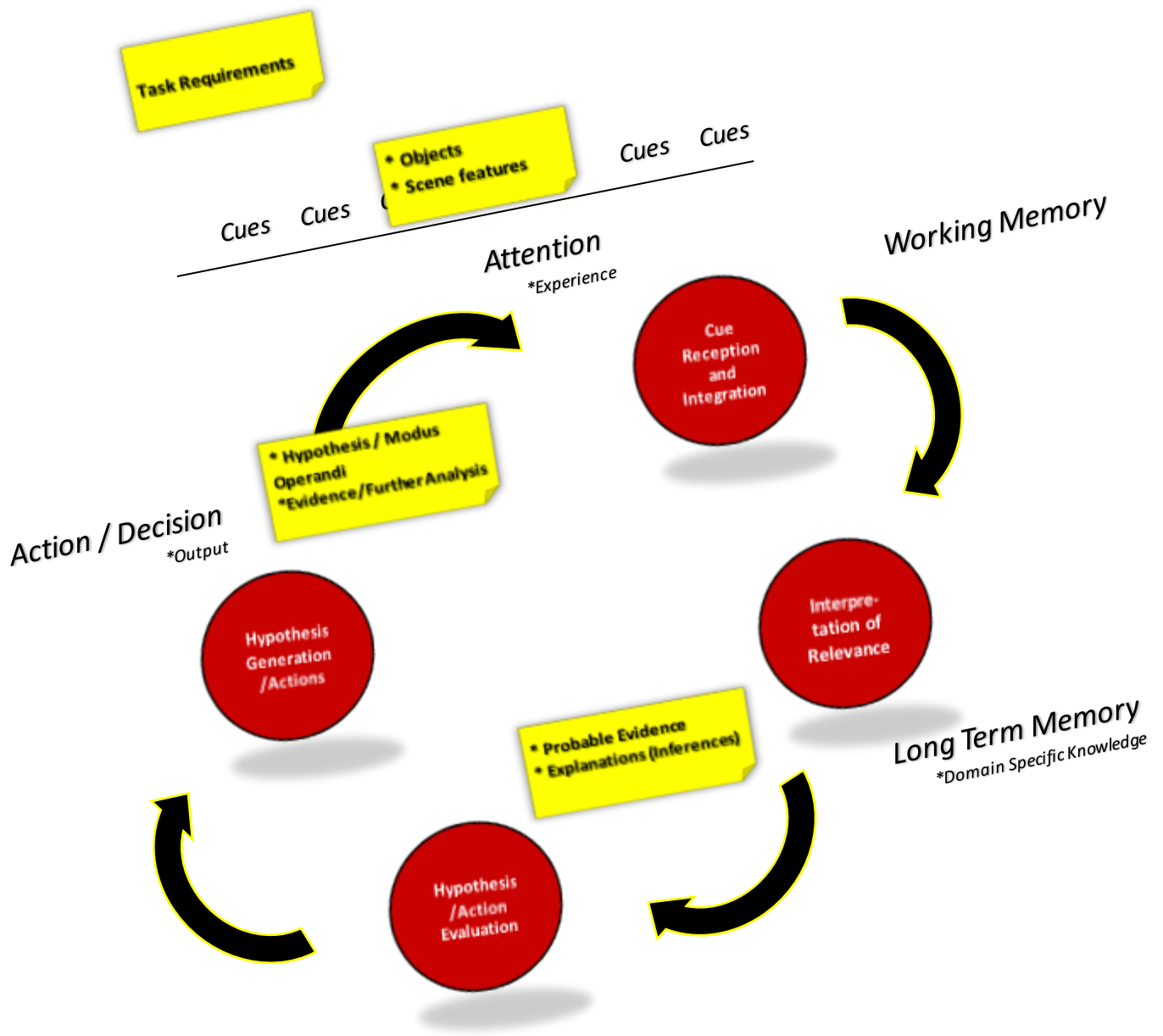


Figure 8.1. The information processing and reasoning process in crime scene examination, showing the involvement of attention, memory and decision making (partly adapted from Wickens, Gordon & Liu, 1998). The diagram shows a course description of the cognitive processes involved in crime scene investigation, where individual processes may be pooled together in larger components.

8.4. Strengths and Weaknesses of the Experiments

The present experiments had several important strengths. First, several eye tracking methods (desktop, mobile) and measures (eye movement characteristics, eye movement patterns, pupil dilation) were used in combination with various non-eye movement measures (e.g., verbal reports, evidence selection, change blindness, memory performance) with the aim of obtaining converging evidence for the role of expertise in crime scene investigation. All these measures are quantitative results, and much less influenced by researchers' interpretation of finding than for example qualitative measures, such as those obtained in focus groups or structured interviews. Second, a range of tasks was used, some based on earlier studies that demonstrated expertise effects in other domains (e.g., change blindness), further probing into converging evidence of expertise effects in crime scene investigation. Third, various levels of expertise were considered, with senior and volume crime scene investigators, first and third year forensic students, and novices, some of whom were age matched to the experts.

There were also some limitations. The main limitation is probably the limited sample size, which could have affected the statistical power to detect differences between experts and novices. Because different levels of expertise effects are difficult to obtain in the same individual, expertise studies tend to rely on between subjects designs. Such designs inherently suffer from the issue that participants tend to differ substantially between each other, which makes it much more difficult to detect significant differences than when comparing different conditions within the same individuals (differences between groups must exceed differences within groups to be able to detect the difference). Sample sizes are often a problem in studies of expertise, as experts are small in numbers (e.g., there are only small number of surgeons in each hospital; Hermens et al., 2013, for a discussion of this problem

in the context of surgery) and often difficult to recruit due to time constraints. For example, in comparisons of expert and novice surgeons, there were 5 experts against 5 novices in Law et al. (2004), 4 experts against 4 novices in Tien et al. (2010) or 8 experts against 6 novices in Wilson et al. (2010). The only way to recruit larger numbers of participants when experts are involved, is by conducting studies across multiple sites. This requires collaboration with many police forces, and it is also possible with research teams across universities. As a possible alternative approach, results may be obtained in smaller studies, and later combined in a meta-analysis. Such an approach, however, requires studies to use very similar approaches, so a method of standardizing the studies would be required. The most obvious solution to the problem of sample size would be to try and increase this in future studies. Experts, however, are often difficult to find, particularly in projects with little funding for participant recruitment and testing across multiple testing sites. If expertise effects are deemed sufficiently important, the restrictions on the sample size are likely to be overcome due to increased availability of funding. Alternative approaches would be to, instead of relying on between subjects designs, on comparisons within the same individual. For simple tasks, such as manipulating objects, this could be achieved by asking participants to conduct the task either with their dominant hand or their non-dominant hand. For more complex tasks, one could rely on training individuals in the task under investigation. For very complex skills, however, training may take a long time, and therefore employing such an approach would rely on the availability of long-term funding. In current climate of relatively short projects with expectations of immediate results, such an approach therefore seems to be less feasible.

A second possible issue is that across the experiments, experts and novices differed in their expertise. These differences make it difficult to compare the results across experiments, as experts in one experiment may be novices in another. This issue was also encountered when

comparing studies on surgical skill (Hermens, Flin & Ahmed, 2013). This limitation is particularly important when one would like to combine studies across studies to increase the statistical power comparing experts and novices, such as in a meta-analysis. Effect sizes in such comparisons may differ because of the difference in tasks, but also because of differences in expertise level differences across studies. The present work particularly suffered from fatigue in the participant group. For the first experiment, participants were keen and were relatively easy to recruit (once the right source was found). After Experiment 1, experts and novices were more difficult to recruit, and this experiment no longer tested the age matched controls that were recruited for Experiment 1. Practical considerations entered the equation in Experiment 3, where the testing site could not easily be moved (being a room inside a church near the police head-quarters). This experiment therefore relied on crime scene investigators only, comparing senior and volume crime scene investigators (novices were difficult to recruit near the police headquarters). For the final experiment, experts were impossible to recruit (due to experiment fatigue and the limited time available for the study), and this experiment therefore relied on forensic and non-forensic students. Differences between selections of experts and novices are found in many studies (e.g., in the domain of surgery, experts in one study may be novices in another). Moreover, the criteria for the “expert” and novice” participants varied across the studies. An alternative approach (as suggested by Chi, 2006) is to follow up novices while they acquire the skills to become experts. Such an approach has the advantages that comparisons between expertise levels are within subjects rather than between subjects, increasing the statistical power to detect differences, and that more fine-grained (depending on the sampling rate) levels of expertise can be studied. However, attractive such an approach, it is not feasible within the time-span of a PhD project (spanning three years, while it may take decades to become a CSI expert). Future studies that can be conducted across a broader time-span, however, should consider a

longitudinal approach. It should also be noted that while experienced CSIs took part in the present experiments, they were perhaps not the top players in their field. Possibly, to study expertise effects, only the highest experts suffice. This would be another domain for future work.

A related issue was that participants were all recruited from a single police department. Reasons were the lack of cooperation of other departments, and limited resources to conduct the study, preventing a broader, multi-site investigation. Sampling from multiple departments would be preferable, however, because it would demonstrate expertise effects independent of the particular department considered, whom, each of them, may have different settings and ways of working.

A further issue may have been that three of the four experiments relied on photographs of (simulated) crime scenes. Photographs have the advantage that the visual stimulus can be kept constant across participants, but they do not provide the rich environment that an actual (simulated) crime scene can offer (with other sensations beyond visual, such as smell, draught, and sounds).

The present work does give an indication as to why it is so difficult to establish criteria for expertise in the context of crime scene investigation. Only recently Kelty et al. (2011) proposed detailed selection criteria and requirements to be part of an A-Team in crime scene investigation. The present results suggest that more research needs to be undertaken into how to define excellence criteria for crime scene investigation. Ideally, objective criteria are sought, that rely on measurable quantities rather than peer judgment, but the present work indicates that establishing these is not going to be straightforward. In the context of border-

personal, some progress has been made in tests to recruit people with outstanding face recognition skills, known as “Super-recognizers” (Robertson et al.,2016), where it was also shown that face recognition could not be easily trained, so selection is key. Conducting similar studies with the specially selected personnel or the collaboration with the police forces which already uses the Keltz et al. (2011) criteria could lead to improved performances of CSIs in general.

8.5. Recommendations for Further Research

8.5.1. Focusing on Team Work

Besides the evaluation of individual process during the tasks, in depth analysis of the organizational or environmental factors or may be the most importantly, the effect of working as part of a team during the task can give crucial insights to provide a basis of understanding on interpersonal processes and aid to develop specific tools for assessing and evaluating performances. Also, often working as a team in serious and major crime scene investigations, apart from the interpersonal factors that affect the process, it can be highly informative to establish the effectiveness of the teams on various examination performances and investigate the coordination, communication and the influence of teamwork. The presence of the other people, even though they do not have any interaction, has been shown to change the viewing behaviours of the participants (Richardson et al., 2012). Therefore, future studies could examine how viewing patterns of CSIs differ while working along compared to as a team. Such studies can give important findings to improve the performances and enhance the teamwork in addition to capture the underlying and may be differed cognitive processes.

8.5.2. Real case work

Due to the dynamic nature of crime scenes which is often subject to organisational constraints in addition to the importance of health hazards to bear in mind, the complex interactions of different factors can be better obtained in real practices instead of the simulated crime scenes. Despite the problems on the control of the conditions in dynamic environments (Chi, 2006), the future studies should consider to collect data in a number of real crime scene settings to improve the basis of understanding the variances in naturalistic environments. Besides the positive contributions of the recent studies of Baber and Butler (2012) and de Gruijter et al. (2016) in controlled simulator environments, additional studies which will be conducted in various natural environments will emphasize the critical aspects of dynamic complex situations and expand the scope of research on the investigation of different factors which have an influence on cognitive processes during the tasks. Therefore, the more naturalistic conditions and practices will lead to exploration of the in-depth analysis of more realistic underlying processes and provide useful insights on the understanding of differences in real cases such as the difference on reasoning or decision making process on real time situations, which was also indicated in Helsloot and Groenendaal (2011, see Chapter 1 for details)

8.5.3. Use eye tracking for feedback and self-reflection

One of the subjects in Wilson-Kovacs (2013) emphasized the lack of feedback on the subsequent processes of the investigation and outcomes. Furthermore, obtaining feedback about the performances in the process has also crucial importance to improve the performance (King, 2009). In the unofficial conversations with the CSIs after the mobile eye tracking study (Chapter 5), the CSIs stated the surprising positive effect of watching their

eye movement recordings and their performances in the study. It is assumed that with respect to the positive responses about the used technology from the participants, the used technology can aid to improve the performances and contribute to learning processes while providing “immediate and informative feedback” about the performances, used strategies and also the errors and failures and corrective actions on the training sessions which can lead to performance improvements (Ericsson & Towne, 2010). In addition to the aforesaid aspects of the development of expertise, the given opportunity of repeated activities will lead to more efficient and effective outcomes which is described as deliberate practice (Ericsson, et al., 1993; Ericsson, 2006; Ericsson & Towne, 2010). Also, in his study King (2009) indicated the importance of self-feedback strategies which would self-monitor the awareness of individuals’ own performances. The guidance of such strategies will foster the rapid developments in the acquisition of expert performance and understanding of the required skills and attitudes to explain expert behaviour in the domains.

Therefore, the use of video recording can provide opportunities to create various training options in practice and lead to obtain more information about the underlying mental processes during the tasks. Thus, the recordings can also be used as part of an experiential learning strategy (e.g. giving instructions or corrective actions from the mentors or coaches (trainers); see King, 2009) while showing various critical incidents to increase and accelerate the expertise in learning and training programs. Throughout the viewing and reviewing the learned and corrected optimal strategies in practice might lead to better performances and outcomes and provide valuable insights about the implicit processes during the tasks at the end (see O’Meara et al., 2015; Browning et al., 2016 for the application of the strategy in healthcare). Also despite the inadequate capabilities of collecting eye movement data which

can give useful data about different levels of perception and cognition, adaptation of cost effective head mounted cameras might also provide important information in the domain.

8.5.4. Additional Tools to Explore the Processes

The neglected nature of studies also revealed that, there is a need to structure a well described assessment tool which can also be applied in the trainings like in the other domains such as surgery (see Yule, Flin, Paterson-Brown & Maran, 2006; Yule et al., 2008). Measuring the performances according to standardized tools will help to boost the performances and scrutinize the representation, elicitation and organization of expert knowledge as well as the nature of tasks and also the task components.

Usually the researchers followed the “task analysis” in various domains for identifying task components and knowledge acquisition and organization for the exploration of human capacity for making decisions and solving problems (Schraagen, 2006). Schraagen (2006) defined task analysis as “what a person is required to do, in terms of actions and/or cognitive processes, to achieve a system goal” (p.185). Especially this analysis provided information to see all required steps of the task, helped to learn what you expect from learners when performing a task, to understand necessary expert knowledge and performance, to develop methods for training, to explore the criteria and to develop tests for the selection of personnel and assessment of performances to see who has the required skills and/or evaluate the required skills in the tasks (Jarodzka, 2010; Scharaagen, 2006). Also with the developments in technology and due to difficulties of task analysis methods applied earlier in tracking and getting insights about cognitive processes, “cognitive task analysis” methods have been developed and employed in the studies instead (Hoffman & Lintern, 2006). The application

of the aforesaid methods in combination with the possible eye tracking while capturing task performance during the overall stages of the crime scene investigation process might give a comprehensive framework for understanding the requirements of the task and also the expertise in the domain.

8.6. Conclusion

The present experiments only revealed small differences between experts' and novices' performance and underlying cognitive and perceptual processes. While relatively small samples were available for the study, the range of measures considered was broad. The experiments employed eye tracking, verbal reporting, selecting evidence, and change detection to assess differences between expert and novice CSIs. If there would have been strong expertise effects, they would have been likely to appear in some if not most of these measures. The relatively weak effects of expertise between experts and novices found here, could pose a general problem in the assessment of expertise in crime scene investigation. This lack of expertise effect and difference may relate to the rather diverse range of perceptual layouts of crime scenes, reducing possible top-down effects of expertise on the deployment of attention (i.e. the crime scene is just a "normal" visual scene without any unique defining visual characteristics, in contrast to say a fingerprint or X-ray image). Although small differences could be important and lead to more efficient and effective outcomes in the overall assessment and examination at the end of the process, it has been very difficult to construct a comprehensive model to explain differences in levels of expertise. The complex nature of the crime scene examination process and the limitations in the present study restrict its overall contribution to understanding the problem. Also, if there are no clear differences in performance or eye movements between experts and novices,

assessment of expertise based on a single test or case is going to be difficult. Instead, overall performance (e.g., the number of cases solved) across a series of cases may need to be considered in order to obtain an informed assessment. However, I believe the wider community of researchers and CSI practitioners will benefit from this research by its future contribution to more useful and practical outcomes to overcome an important real world problem. Despite the limitations and difficulties, mobile eye tracking during actual crime scene investigation provides the most promising method for developing assessment and training tools on the basis of eye movements in crime scene investigation.

List of References

Adderley, R., Smith, L. L., Bond, J. W., & Smith, M. (2012). Physiological measurement of crime scene investigator stress. *International Journal of Police Science & Management*, *14*(2), 166-176.

Afonso, J., Garganta, J., Mcrobert, A., Williams, A. M., & Mesquita, I. (2012). The perceptual cognitive processes underpinning skilled performance in volleyball: Evidence from eye-movements and verbal reports of thinking involving an in situ representative task. *Journal of Sports Science & Medicine*, *11*(2), 339–345.

Ahlstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, *36*(7), 623–636. doi:10.1016/j.ergon.2006.04.002

Ahmidi, N., Hager, G. D., Ishii, L., Fichtinger, G., Gallia, G. L., & Ishii, M. (2010). Surgical task and skill classification from eye tracking and tool motion in minimally invasive surgery. *Medical Image Computing and Computer-Assisted Intervention: MICCAI. International Conference on Medical Image Computing and Computer-Assisted Intervention*, *13*(Pt 3), 295-302.

Ahmidi, N., Ishii, M., Fichtinger, G., Gallia, G. L., & Hager, G. D. (2012). An objective and automated method for assessing surgical skill in endoscopic sinus surgery using eye-tracking and tool-motion data. *International Forum of Allergy and Rhinology*, 2(6), 507-515. doi:10.1002/alr.21053

Allard, E. S., Wadlinger, H. A., & Isaacowitz, D. M. (2010). Positive gaze preferences in older adults: assessing the role of cognitive effort with pupil dilation. *Aging, Neuropsychology, and Cognition*, 17(3), 296-311. doi:10.1080/13825580903265681

Allison, L. & Adams, K. (2012) *Visual Searching in Forensic Experts*. Poster presented at Tobii Eye Tracking Conference on Behavioural Research, Belgium

Anderson, J. R., Bothell, D., & Douglass, S. (2004). Eye Movements Do Not Reflect Retrieval Processes Limits of the Eye-Mind Hypothesis. *Psychological Science*, 15(4), 225-231. doi:10.1111/j.0956-7976.2004.00656.x

Anderson, N. C., Bischof, W. F., Laidlaw, K. E. W., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior Research Methods*, 45(3), 842-856. doi:10.3758/s13428-012-0299-5

Angie, A. D., Connelly, S., Waples, E. P., & Kligyte, V. (2011). The influence of discrete emotions on judgement and decision-making: A meta-analytic review. *Cognition and Emotion*, 25(8), 1393-1422. doi:10.1080/02699931.2010.550751

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology, 10*(3), 229-240. doi:10.1037/1089-2680.10.3.229

Atkins, M. S., Jiang, X., Tien, G., & Zheng, B. (2012). Saccadic delays on targets while watching videos. *Eye Tracking Research and Applications Symposium (ETRA)*, 405-408. doi:10.1145/2168556.2168648

Baber, C. (2010). Distributed cognition at the crime scene. *AI and Society, 25*(4), 423-432. doi:10.1007/s00146-010-0274-6

Baber, C. & Butler, M. (2012). Expertise in crime scene examination: comparing search strategies of expert and novice crime scene examiners in simulated crime scenes. *Human Factors, 54*(3), 413-24.

Ballard, D. H. & Hayhoe, M. M. (2009). Modelling the role of task in the control of gaze. *Visual Cognition, 17*(6-7), 1185-1204. doi:10.1080/13506280902978477

Balslev, T., Jarodzka, H., Holmqvist, K., De Grave, W., Muijtjens, A. M., Eika, B., Van Merriënboer, J. & Scherpbier, A.J.J.A. (2012) Visual expertise in paediatric neurology. *European Journal of Paediatric Neurology, 16*(2) 161-166.

Beard, D. V., Johnston, R. E., Toki, O., & Wilcox, C. (1990). A study of radiologists viewing multiple computed tomography examinations using an eyetracking device. *Journal of Digital Imaging, 3*(4), 230-237. doi:10.1007/BF03168120

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91(2), 276-292. doi:10.1037/0033-2909.91.2.276

Beatty, J. & Lucero-Wagoner, B. (2000). Pupillary System. In J.T. Cacioppo, L.G. Tassinary, & G. Berntson (Eds.). *Handbook of Psychophysiology* (pp. 142–162). Cambridge University Press.

Beck, M.R. & Levin, D.T. (2003). The role of representational volatility in recognizing pre- and postchange objects. *Perception and Psychophysics*, 65(3) 458-468.

Beck, M.R., Martin, B.A., Smitherman, E., & Gaschen, L. (2013). Eyes-on training and radiological expertise: An examination of expertise development and its effects on visual working memory. *Human Factors*, 55(4) 747-763.

Beck, M. R., Peterson, M. S., & Angelone, B. L. (2007). The roles of encoding, retrieval, and awareness. *Memory & Cognition*, 35(4), 610-620.

Beck, M. R., Levin, D. T., & Angelone, B. (2007). Change blindness blindness: Beliefs about the roles of intention and scene complexity in change detection. *Consciousness and Cognition*, 16(1), 31-51.

Becker, S.I. (2014). Guidance of Attention by Feature Relationships: The End of the Road for Feature Map Theories? In: M. Horsley, M. Elliot, M. A. Knight & R. Reilly (Eds.) *Current Trends in Eye Tracking Research* (pp. 37-50). Switzerland: Springer International Publishing

Bellenkes, A. H., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine*, 68(7), 569–579

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., & Montanari, R. (2011). Driver workload and eye blink duration. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14(3), 199-208. doi:10.1016/j.trf.2010.12.001

Benson, P. J., Beedie, S. A., Shephard, E., Giegling, I., Rujescu, D., & St. Clair, D. (2012). Simple viewing tests can detect eye movement abnormalities that distinguish schizophrenia cases from controls with exceptional accuracy. *Biological Psychiatry*, 72(9), 716-724. doi:10.1016/j.biopsych.2012.04.019

Bentivoglio, A. R., Bressman, S. B., Cassetta, E., Carretta, D., Tonali, P., & Albanese, A. (1997). Analysis of blink rate patterns in normal subjects. *Movement Disorders*, 12(6), 1028-1034. doi:10.1002/mds.870120629

Biederman, I. (2000). Recognizing depth-rotated objects: A review of recent research and theory. *Spatial Vision*, 13(2-3), 241-253. doi:10.1163/156856800741063

Birmingham, E., Bischof, W., & Kingstone, A. (2008a). Gaze selection in complex social scenes. *Visual Cognition*, *16*(2-3), 341-355. doi:10.1080/13506280701434532

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008b). Social attention and real-world scenes: The roles of action, competition and social content. *Quarterly Journal of Experimental Psychology*, *61*(7), 986-998. doi:10.1080/17470210701410375

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Get real! resolving the debate about equivalent social stimuli. *Visual Cognition*, *17*(6-7), 904-924. doi:10.1080/13506280902758044

Block, R.A.(2009). Intent to remember briefly presented human faces and other pictorial stimuli enhances recognition memory. *Memory & Cognition*, *37*(5), 667-678. doi:10.3758/MC.37.5.667

Bojko, A.A. (2009). Informative or misleading? Heatmaps deconstructed. In *Human-computer interaction. New trends* (pp. 30-39). Springer Berlin Heidelberg. doi:10.1007/978-3-642-02574-7_4

Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: Sex differences in picture processing. *Emotion*, *1*(3), 300-319. doi:10.1037/1528-3542.1.3.300

Bradley, M.M., Houbova, P., Miccoli, L., Costa, V.C., & Lang, P.J. (2011). Scan patterns when viewing natural scenes: Emotion, complexity, and repetition. *Psychophysiology*, 48(11): 1544-1553. doi:10.1111/j.1469-8986.2011.01223.x

Bradley, M. M., Miccoli, L., Escrig, M. A. & Lang P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45(4), 602–607. doi:10.1111/j.1469-8986.2008.00654.x

Broadbent, M., Horsley, M., Birks, M. & Persaud, N. (2014). Comparing Novice and Expert Nurses in Analysing Electrocardiographs (ECGs) Containing Critical Diagnostic Information: An Eye Tracking Study of the Development of Complex Nursing Visual Cognitive Skills. In: M. Horsley, M. Elliot, M. A. Knight & R. Reilly (eds.) *Current Trends in Eye Tracking Research*. (pp. 297-316.) Switzerland: Springer International Publishing.

Brosch, T., Scherer, K., Grandjean, D., & Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly*, 143: 1–10. doi: 10.4414/smw.2013.13786.

Browning, M., Cooper, S., Cant, R., Sparkes, L., Bogossian, F., Williams, B., O'Meara, P., Ross, L., Munro, G., & Black, B. (2016) The use and limits of eye-tracking in high-fidelity clinical scenarios: A pilot study. *International Emergency Nursing*, 25, 43–47. doi:10.1016/j.ienj.2015.08.002

Buchanan, T. W. (2007). Retrieval of Emotional Memories. *Psychological Bulletin*, 133(5), 761–779. <http://doi.org/10.1037/0033-2909.133.5.761>

Buckee, C., Kneller, B., & Peakall, J. (2001). Turbulence structure in steady solute-driven gravity currents. In: W.D. McCaffrey, B. Kneller, & J. Peakall (Eds.). *Particulate Gravity Currents. Special Publication of the International Association of Sedimentology*. (pp. 173–188) S.P. 31. Blackwell Science, Oxford,

Busey, T., Yu, C., Wyatte, D. & Vanderkolk, J. (2013) Temporal sequences quantify the contributions of individual fixations in complex perceptual matching tasks. *Cognitive Science*, 37(4) 731-756. doi:10.1111/cogs.12029

Busey, T., Yu, C., Wyatte, D., Vanderkolk, J., Parada, F. & Akavipat, R. (2011) Consistency and variability among latent print examiners as revealed by eye tracking methodologies. *Journal of Forensic Identification*, 61(1) 60-91.

Buswell, G.T. (1935). *How people look at pictures: A Study of the Psychology of Perception in Art*. Chicago: University of Chicago Press.

Caldara, R., & Miellet, S. (2011). iMap: A novel method for statistical fixation mapping of eye movement data. *Behavior Research Methods*, 43(3), 864-878. doi:10.3758/s13428-011-0092-x

Calvo, M. G., & Avero, P. (2005). Time course of attentional bias to emotional scenes in anxiety: Gaze direction and duration. *Cognition & Emotion*, 19(3), 433-451.

Cardinal, R.N. & Aitken, M.R. (2006). *ANOVA for the Behavioural Sciences Researcher*. Lawrence Erlbaum Associates, Inc. Publishers, New Jersey.

Carmi, R. & Itti, L. (2006) Visual causes versus correlates of attentional selection in dynamic scenes. *Vision Research*, 46(26) 4333-4345.

Carmichael, A., Larson, A., Gire, E., Loschky, L. & Rebello, N.S. (2010). How does visual attention differ between experts and novices on physics problems? *AIP Conference Proceedings*, 1289(1), 93-96. doi:10.1063/1.3515257

Carniglia, E., Caputi, M., Manfredi, V., Zambarbieri, D., & Pessa, E. (2012). The influence of emotional picture thematic content on exploratory eye movements. *Journal of Eye Movement Research*, 5(4), 1-9.

Carpenter, R. H. S. (1988). *Movements of the Eyes (2nd ed.)*. London: Pion.

Carrasco, M. (2011) Visual attention: The past 25 years. *Vision Research*, 51(13) 1484-1525.

Carson, D. (2009). The abduction of Sherlock Holmes. *International Journal of Police Science & Management*, 11 (2), 193-202

Castro, S. L., & César, F. L. (2014). Age and Musical Expertise Influence Emotion Recognition in Music. *Music Perception*, 32(2), 125–142. doi:10.1525/MP.2014.32.2.125

Chainay, H., Michael, G. A., Vert-pré, M., Landré, L., & Plasson, A. (2012). Emotional enhancement of immediate memory: Positive pictorial stimuli are better recognized than neutral or negative pictorial stimuli. *Advances in Cognitive Psychology*, 8(3), 255-266. doi: 10.5709/acp-0121-1

Champod, C. (2014) Research focused mainly on bias will paralyse forensic science. *Science and Justice*, 54(2) 107-109.

Chapman, P. R. & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, 27(8), 951-964.

Charness, N., Reingold, E.M., Pomplun, M. & Stampe, D.M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory and Cognition*, 29(8) 1146-1152.

Charness, N. & Tuffiash, M. (2008). The role of expertise research and human factors in capturing, explaining, and producing superior performance. *Human Factors*, 50(3), 427-432.

Chase, W. G. & Simon, H. A. (1973) The mind's eye in chess. In: W. G. Chase (Ed.), (pp.215-281) New York: Academic Press.

Chi, M. T. H. (2006). Laboratory methods for assessing experts' and novices' knowledge. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.). *The Cambridge Handbook of Expertise and Expert Performance*, (pp. 167-184). New York: Cambridge University Press.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P.J. Feltovich, & R.R. Hoffman (Eds.) *The Cambridge Handbook of Expertise and Expert Performance*, (pp. 21-30). New York: Cambridge University Press

Chi, M.T.H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121-152.

Chipchase, S.Y. & Chapman, P. (2013). Trade-offs in visual attention and the enhancement of memory specificity for positive and negative emotional stimuli. *The Quarterly Journal of Experimental Psychology*, 66(2), 277-298. doi:10.1080/17470218.2012.707664

Chisholm, S. L., Caird, J. K., Lockhart, J. A., Teteris, L. E., & Smiley, A. (2006). Novice and experienced driving performance with cell phones. *Proceedings of the Human Factors and Ergonomics Society*, 2354-2358.

Christianson, S. A. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, 112(2), 284. doi:10.1037/0033-2909.112.2.284

Coco, M. I. (2009). The statistical challenge of scan-path analysis. *Proceedings - 2009 2nd Conference on Human System Interactions, HSI'09*, 372-375. doi: 10.1109/HSI.2009.5091008

Cole, S. (2013). Forensic Science Reform: Out of the Laboratory and into the Crime Scene. *Texas Law Review*, 91, 123-136.

Collewijn, H., & Kowler, E. (2008). The significance of microsaccades for vision and oculomotor control. *Journal of Vision*, 8(14:20), 1-21. doi:10.1167/8.14.20

Collewijn, H., Van der Mark, F., & Jansen, T. C. (1975). Precise recording of human eye movements. *Vision Research*, 15(3), 447-450. doi:10.1016/0042-6989(75)90098-X

Cooke, L., & Cuddihy, E. (2005). Using eye tracking to address limitations in think-aloud protocol. *Proceedings of International Professional Communication Conference*, 653-658.

Cooper, L., Gale, A., Darker, I., Toms, A., & Saada, J. (2009). Radiology image perception and observer performance: How does expertise and clinical information alter interpretation? Stroke detection explored through eye-tracking. *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, Lake Buena Vista, FL, USA. doi:10.1117/12.811098

Corcoran, S. A. (1986). Planning by expert and novice nurses in cases of varying complexity. *Research in Nursing & Health*, 9(2), 155-162.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, 34(4), 613-617.

Crane H.D. & Steele, C.M. (1985). Generation-V dual-Purkinje-Image eye tracker. *Applied Optics*, 24 (4), 527-537

Crispino, F. (2008) Nature and place of crime scene management within forensic sciences. *Science and Justice*, 48(1) 24-28.

Cristino, F., Mathôt, S., Theeuwes, J., & Gilchrist, I. D. (2010). ScanMatch: A novel method for comparing fixation sequences. *Behavior Research Methods*, 42(3), 692-700. doi:10.3758/BRM.42.3.692

Crundall, D. (2016). Hazard prediction discriminates between novice and experienced drivers. *Accident Analysis and Prevention*. 86, 47-58. doi: 10.1016/j.aap.2015.10.006

Crundall, D., Chapman, P., Trawley, S., Collins, L., Van Loon, E., Andrews, B., & Underwood, G. (2012). Some hazards are more attractive than others: Drivers of varying experience respond differently to different types of hazard. *Accident Analysis and Prevention*, 45, 600-609. doi:10.1016/j.aap.2011.09.049

Crundall, D., Underwood, G., & Chapman, P. (1999). Driving experience and the functional field of view. *Perception*, 28(9), 1075-1087.

Davies, G., & Hine, S. (2007). Change blindness and eyewitness testimony. *The Journal of psychology: Interdisciplinary and Applied*, 141(4), 423-434. doi:10.3200/JRLP.141.4.423-434

De Bie, J. (1985). An afterimage vernier method for assessing the precision of eye movement monitors: Results for the scleral coil technique. *Vision Research*, 25(9), 1341-1343. doi:10.1016/0042-6989(85)90051-3

De Groot, A.D. (1965). *Thought and Choice in Chess*. The Hague: Mouton.

De Gruijter, M., De Poot C. J. & Elffers, H. (2016) The Influence of New Technologies on the Visual Attention of CSIs Performing a Crime Scene Investigation. *Journal of Forensic Sciences*, 61(1), 43-51. doi: 10.1111/1556-4029.12904

DeLaurentis, P. (2009). 3D Scanning: A New Tool for Cracking Tough Cases. *Forensic Magazine*, 6(1), 37-40.

Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R., & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavior Research Methods*, 44(4), 1079-1100. doi:10.3758/s13428-012-0212-2

Dell'Osso, L. & Daroff, R. (1999). Eye Movement Characteristics and Recording Techniques. In J. S. Glaser, (Ed.). *Neuro-ophthalmology* (3rd ed.), (pp. 327-344). Lippincott Williams & Wilkins, Philadelphia,

Denzin, N. K., & Lincoln, Y. S. (1994). *Handbook of qualitative research*. Sage Publications, Inc.

Di Stasi, L.L., Contreras, D., Cándido, A., Canas, J. & Catena, A. (2011). Behavioral and eyemovement measures to track improvements in driving skills of vulnerable road users: first-time motorcycle riders. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14(1), 26–35.

Dicks, M., Button, C., & Davids, K. (2010). Examination of gaze behaviors under in situ and video simulation task constraints reveals differences in information pickup for perception and action. *Attention, Perception, and Psychophysics*, 72(3), 706-720. doi:10.3758/APP.72.3.706

Dietz, J., Bradley, M.M., Okunc, M.S., & Bowersa, D. (2011). Emotion and ocular responses in Parkinson's Disease, *Neuropsychologia*, 49(12): 3247–3253. doi:10.1016/j.neuropsychologia. 2011.07.029

Donovan, T., Manning, D.J., & Crawford, T. (2008) Performance changes in lung nodule detection following perceptual feedback of eye movements. *Progress in Biomedical Optics and Imaging - Proceedings of SPIE - The International Society for Optical Engineering*, 6917. doi: 10.1117/12.768503

Drew, T., Võ, M. L., & Wolfe, J. M. (2013). The invisible gorilla strikes again: Sustained inattention blindness in expert observers. *Psychological Science*, 24(9), 1848-1853. doi:10.1177/0956797613479386

Dror, I. E. (2008). Biased brains. *Police Review*, 116, 20-23.

Dror, I. E. (2011). The paradox of human expertise: Why experts get it wrong. *The Paradoxical Brain* (pp. 177-188). doi:10.1017/CBO9780511978098.011

Dror, I. E. (2015). Cognitive neuroscience in forensic science: Understanding and utilizing the human element. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1674). doi:10.1098/rstb.2014.0255

Dror, I. E., & Charlton, D. (2006). Why experts make errors. *Journal of Forensic Identification*, 56(4), 600-616.

Dror, I. E. & Cole, S. A. (2010). The vision in blind justice: Expert perception, judgement, and visual cognition in forensic pattern recognition. *Psychonomic Bulletin & Review*, 17(2), 161–167. doi:10.3758/PBR.17.2.161

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice (2nd ed.)*. London: Springer.

Duchowski, A. T., Driver, J., Jolaoso, S., Tan, W., Ramey, B. N., & Robbins, A. (2010). Scanpath comparison revisited. *Eye Tracking Research and Applications Symposium (ETRA)*, 219-226. doi:10.1145/1743666.1743719

Dyer, A.G., Found, B., & Rogers, D. (2006) Visual attention and expertise for forensic signature analysis. *Journal of Forensic Sciences*, 51(6) 1397-1404.

Dyer, A.G., Found, B., & Rogers, D. (2008) An insight into forensic document examiner expertise for discriminating between forged and disguised signatures. *Journal of Forensic Sciences*, 53(5) 1154-1159.

Eccleston, P., Werneke, U., Armon, K., Stephenson, T., & MacFaul, R. (2001). Accounting for overlap? an application of mezzich's statistic to test interrater reliability of interview data on parental accident and emergency attendance. *Journal of Advanced Nursing*, 33(6), 784-790. doi:10.1046/j.1365-2648.2001.01718.x

Eger, N., Ball, L.J., Stephens, R. & Dodd, J. (2007). Cueing Retrospective Verbal Reports in Usability Testing Through Eye-Movement Replay. *In Proceedings of BCS HCI 2007*.

Einhäuser, W., Stout, J., Koch, C. & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences*, 105, 1704–1709. doi:10.1073/pnas.0707727105

Elling, S., Lentz, L. & de Jong, M. (2011). Retrospective think-aloud method: Using eye movements as an extra cue for participants' verbalizations. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1161-1170.

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: a critical review. In M.R. Endsley & D.J. Garland (Eds.). *Situation awareness analysis and measurement*. (pp. 3–32) Lawrence Erlbaum, Mahwah, NJ.

- Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, 43(9), 1035-1045. doi:10.1016/S0042-6989(03)00084-1
- Engbert, R., & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences*, 103(18), 7192-7197. doi:10.1073/pnas.0509557103
- Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological Review*, 112(4), 777-813. doi:10.1037/0033-295X.112.4.777
- Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks, In K. A. Ericsson, N. Charness, P. Feltovich & R. R. Hoffman (Eds.) *The Cambridge handbook of expertise and expert performance*, (pp.223–241). Cambridge University Press, New York.
- Ericsson, K. A. & Kintsch, W. (1995) Long-term working memory. *Psychological Review*, 102(2) 211-245.
- Ericsson, K.A., Krampe, R.T. & Tesch-Römer, C. (1993) The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3) 363-406.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: Bradford books/MIT Press.

Ericsson, K. A. & Simon, H.A. (1993) *Protocol Analysis: Verbal reports as data. Revised Edition*. Cambridge, Massachusetts: The MIT Press.

Ericsson, K. A. & Smith, J. (1991). Prospects and limits in the empirical study of expertise: an introduction. In K.A. Ericsson & J. Smith, (Eds.) *Toward a General Theory of Expertise: Prospects and Limits*. (pp.1–38). Cambridge: Cambridge University Press

Ericsson, K. A. & Towne, T.J. (2010). Expertise. *Cognitive Science*, 1(3), 404–416.
doi:10.1002/wcs.47

Evans, K. K., Cohen, M. A., Tambouret, R., Horowitz, T., Kreindel, E., & Wolfe, J. M. (2011). Does visual expertise improve visual recognition memory?. *Attention, Perception & Psychophysics*, 73(1), 30–35. <http://doi.org/10.3758/s13414-010-0022-5>

Evans, K. M., Jacobs, R. A., Tarduno, J. A., & Pelz, J. B. (2012). Collecting and Analyzing Mobile Eye-tracking Data in Outdoor Environments. *Journal of Eye Movement Research* 5(2):6, 1-19

Evans, K., Marom, E., Godoy, M., Palacio, D., Sagebiel, T., Cuellar, S.B., McEntee, M., Brennan, P., Tian, C. and Haygood, T. (2016) Radiologists remember mountains better than radiographs, or do they? *Journal of Medical Imaging*, 3(1), 011005; <http://dx.doi.org/10.1117/1.JMI.3.1.011005>.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Fawcett, J. M., Russell, E. J., Peace, K. A., & Christie, J. (2013). Of guns and geese: a meta-analytic review of the “weapon focus” literature. *Psychology, Crime & Law*, *19*(1), 35-66.

Feil, A. & Mestre, J.P. (2010): Change Blindness as a Means of Studying Expertise in Physics, *Journal of the Learning Sciences*, *19*(4), 480-505.

Feltovich, P. J., Prietula, M. J., & Ericsson, K. A. (2006). Studies of Expertise from Psychological Perspectives. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.). *The Cambridge Handbook of Expertise and Expert Performance*, (pp.41-67). New York: Cambridge University Press

Ferman, L., Collewijn, H., Jansen, T. C., & Van den Berg, A. V. (1987). Human gaze stability in the horizontal, vertical and torsional direction during voluntary head movements, evaluated with a three-dimensional scleral induction coil technique. *Vision Research*, *27*(5), 811-828. doi:10.1016/0042-6989(87)90078-2

Fiedler, S., & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology*, *3*(OCT). doi:10.3389/fpsyg.2012.00335

Field, A. P. (2009). *Discovering statistics using SPSS (3rd ed.)*. London: SAGE.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, 22(04), 661-674. doi:10.1017/S0140525X99002150

Fischer, H., Sandblom, J., Gavazzeni, J., Fransson, P., Wright, C.I., & Bäckman, L. (2005). Age-differential patterns of brain activation during perception of angry faces. *Neuroscience Letters*, 386(2), 99-104. doi:10.1016/j.neulet.2005.06.002

Fisher, B. A., & Fisher, D. R. (2012). *Techniques of crime scene investigation*. CRC Press.

Fitts, P. M., Jones, R. E., & Milton, J. L. (2005). Eye movements of aircraft pilots during instrument-landing approaches. In N. Moray (Ed.) *Ergonomics: Psychological Mechanisms and Models in Ergonomics* (pp.56-66). New York: Taylor & Francis.

Flin, R., Youngson, G., & Yule, S. (2007). How do surgeons make intraoperative decisions?. *Quality and Safety in Health Care*, 16(3), 235-239. doi:10.1136/qshc.2006.020743

Flowe, H. & Cottrell, G.W. (2011) An examination of simultaneous lineup identification decision processes using eye tracking. *Applied Cognitive Psychology*, 25(3) 443-451.

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 and Performance*, 24(3) 847-858.

Foulsham, T., Chapman, C., Nasiopoulos, E., & Kingstone, A. (2014). Top-down and bottom-up aspects of active search in a real-world environment. *Canadian Journal of Experimental Psychology*, 68(1), 8–19. doi: 10.1037/cep0000004

Foulsham, T. & Underwood, G. (2007) How does the purpose of inspection influence the potency of visual salience in scene perception? *Perception*, 36(8) 1123-1138.

Foulsham T., Walker E. & Kingstone A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, 51(17), 1920–1931, doi:10.1016/j.visres.2011.07.002.

Found, B. (2015). Deciphering the human condition: the rise of cognitive forensics. *Australian Journal of Forensic Sciences*, 47(4), 386-401. doi: 10.1080/00450618.2014.965204

Fox, J., Merwin, D., Marsh, R., McConkie, G., & Kramer, A. (1996, October). Information extraction during instrument flight: An evaluation of the validity of the eye-mind hypothesis. *Proceedings of the Human Factors and Ergonomics Society*, 77-81.

Freeman, J.B. & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, 42(1), 226-241. doi:10.3758/BRM.42.1.226

Freeman, J.B., Dale, R., & Farmer, T.A. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*, 2, 59. doi:10.3389/fpsyg.2011.00059

Friesen, C. K., & Kingstone, A. (2003). Covert and overt orienting to gaze direction cues and the effects of fixation offset. *Neuroreport*, *14*(3), 489-493.

Galpin, A., Underwood, G., & Crundall, D. (2009). Change blindness in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(2), 179-185. doi:10.1016/j.trf.2008.11.002

Gaspar, J.G., Neider, M.B., Simons, D.J., McCarley, J.S., & Kramer, A.F. (2013). Change detection: training and transfer. *PLoS One*, *8*(6), 1-7. doi: 10.1371/journal.pone.0067781

Gavazzeni, J., Wiens, S., & Fischer, H. (2008). Age effects to negative arousal differ for self-report and electrodermal activity. *Psychophysiology*, *45*(1), 148-151. doi:10.1111/j.1469-8986.2007.00596.x

Geangu, E., Hauf, P., Bhardwaj, R. & Bentz, W. (2011) Infant Pupil Diameter Changes in Response to Others' Positive and Negative Emotions. *PLoS One* *6*(11), e27132. doi:10.1371/journal.pone.0027132

Gegenfurtner, A., Lehtinen, E. & Säljö, R. (2011) Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, *23*(4) 523-552.

Gerjets, P., Kammerer, Y. & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. *Learning and Instruction, 21*(2), 220-231.

Gibson, E. J., & Levin, H. (1975). *The psychology of reading*. The MIT press.

Gidlöf, K., Wallin, A., Dewhurst, R., & Holmqvist, K. (2013). Using eye tracking to trace a cognitive process: Gaze behaviour during decision making in a natural environment. *Journal of Eye Movement Research, 6*(1), 1-14.

Gilbert, A. C., Boucher, V. J., & Jemel, B. (2014). Perceptual chunking and its effect on memory in speech processing: ERP and behavioral evidence. *Frontiers in Psychology, 5*, 220. doi:10.3389/fpsyg.2014.00220

Gingras, B., Marin M.M., Puig-Waldmüller, E. & Fitch, W.T. (2015) The Eye is Listening: Music-Induced Arousal and Individual Differences Predict Pupillary Responses. *Frontiers: Human Neuroscience, 9*:619. doi:10.3389/fnhum.2015.00619

Glaser, E., Mendrek, A., Germain, M., Lakis, N., & Lavoie, M. E.(2012). Sex differences in memory of emotional images: A behavioral and electrophysiological investigation. *International Journal of Psychophysiology, 85*(1),17-26. doi:10.1016/j.ijpsycho.2012.01.007

Glöckner, A., & Herbold, A. K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, 24(1), 71-98.

Halpern, D. F., & Wai, J. (2007). The world of competitive Scrabble: Novice and expert differences in visuospacial and verbal abilities. *Journal of Experimental Psychology: Applied*, 13(2), 79-94. doi:10.1037/1076-898X.13.2.79

Harle, S. K., & Vickers, J. N. (2001). Training quiet eye improves accuracy in the basketball free throw. *Sport Psychologist*, 15(3), 289-305.

Harrison, K. (2006). Is crime scene examination science, and does it matter anyway? *Science & Justice*, 46(2), 65-68.

Hayhoe, M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, 3(1), 49-63. doi:10.1167/3.1.6

Hegarty, M., Canham, M.S. and Fabrikant, S.I. (2010) Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning Memory and Cognition*, 36(1) 37-53.

Hegarty, M., Mayer, R., & Monk, C. (1995). Comprehension of arithmetic word problems: a comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology*, 87(1), 18-32

Helman, E., Stoller, R.M., & Freeman, J.B. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. *Group Processes and Intergroup Relations*, 18(3),384-401.

Helsloot, I. & Groenendaal, J. (2011) Naturalistic decision making in forensic science: toward a better understanding of decision making by forensic team leaders. *Journal of Forensic Sciences*, 56(4), 890-897. doi: 10.1111/j.1556-4029.2011.01714

Heminghaus, J. & Duchowski, A. T. (2006). *Icomp: A tool for scanpath visualization and comparison*. Poster presented at the meeting of SIGGRAPH 2006
doi:10.1145/1179622.1179836

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7(11), 498-504.

Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science*, 16(4), 219-222.

Henderson, J.M., Brockmole, J.R., Castelano, M.S. & Mack, M. (2007) Visual saliency does not account for eye movements during visual search in real-world scenes. In R. P. G van Gompel, M. H. Fischer, W. S. Murray & R. L Hill (Eds.) *Eye movements: A window on mind and brain* (pp.537-562). Oxford, UK: Elsevier.

Henderson, J.M. & Hollingworth, A. (1999) The role of fixation position in detecting scene changes across saccades. *Psychological Science*, 10(5), 438-443.

- Henderson, J.M. & Hollingworth, A. (2003) Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception and Psychophysics*, 65(1) 58-71.
- Hermens, F., Flin, R. & Ahmed, I. (2013). Eye movements in surgery: A literature review. *Journal of Eye Movement Research*, 6(4), 1-11.
- Ho, S., Foulsham, T., & Kingstone, A. (2015). Speaking and listening with the eyes: gaze signaling during dyadic interactions. *PloS One*, 10(8), e0136905. doi:10.1371/journal.pone.0136905
- Hodgson, T.L., Bajwa, A., Owen, A.M. & Kennard, C. (2000). The strategic control of gaze direction in the Tower of London task . *Journal of Cognitive Neuroscience*, 12(5), 894–907.
- Hoffman, R. R. & Lintern, G. (2006). Eliciting and Representing the Knowledge of Experts, In K. A. Ericsson, N. Charness, P. Feltovich & R. R. Hoffman (Eds.) *The Cambridge Handbook of Expertise and Expert Performance* (pp. 203-222). Cambridge University Press, New York.
- Hoffman, K. A., Aitken, L. M., & Duffield, C. (2009). A comparison of novice and expert nurses' cue collection during clinical decision-making: Verbal protocol analysis. *International Journal of Nursing Studies*, 46(10), 1335-1344.
- Hollingworth, A. & Henderson, J.M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Visual Cognition*, 7(1-3), 213-235.

Holmqvist, K., Nystrom, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Hope, L. & Wright, D. (2007) Beyond unusual? examining the role of attention in the weapon focus effect. *Applied Cognitive Psychology*, 21(7) 951-961.

Horswell, J. (2004). *The practice of crime scene investigation*. CRC Press.

Hosking, S.G., Liu, C.C. & Bayly M. (2010) The visual search patterns and hazard responses of experienced and inexperienced motorcycle riders. *Accident Analysis and Prevention*, 42(1),196–202. doi:10.1016/j.aap.2009.07.023

Houck, M.M. (2015) *Professional Issues in Forensic Sciences*. Oxford: Academic Press.

Hulse, L. M. & Memon, A. (2006). Fatal impact? The effects of emotional arousal and weapon presence on police officers' memories for a simulated crime. *Legal and Criminological Psychology*, 11(2), 313–325. doi:10.1348/135532505X58062

Humphrey, K., Underwood, G., & Lambert, T. (2012). Salience of the lambs: a test of the salience map hypothesis with pictures of emotive objects. *Journal of Vision*, 12(1), 1–15. doi: 10.1167/12.1.22

Humphreys, L., Underwood, G. & Chapman, P. (2010). Enhanced memory for emotional pictures: A product of increased attention to affective stimuli? *European Journal of Cognitive Psychology*, 22(8), 1235-1247. doi:10.1080/09541440903427487

Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction*, 20(2), 172–176. doi:10.1016/j.learninstruc.2009.02.013

Hyrskykari, A., Ovaska, S., Majaranta, P., Räihä, K.-J. & Lehtinen, M. (2008). Gaze Path Stimulation in Retrospective Think Aloud. *Journal of Eye Movement Research*. 2(4):5, 1-18

Hyun, S. H., Kim, K. T., Song, S. M., Yoon, J. S., Lee, J. I., Cho, S. M., & Sohn, Y. W. (2006). Effects of expertise and situation complexity on visual attention and action planning for air traffic control. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 420–423), San Francisco, USA.

Inman, K. & Rudin, N. (2001). *Principles and practice of criminalistics: The profession of forensic science*. CRC Press

Innes, M., (2003). *Investigating murder: detective work and the police response to criminal homicide*. Oxford: Oxford University Press.

Irwin, D. E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology*, 36(1), 1-27.

Irwin, D. E., Carlson-Radvansky, L. A., & Andrews, R. V. (1995). Information processing during saccadic eye movements. *Acta Psychologica*, *90*(1), 261-273.

Irwin, D. E., & Zelinsky, G. J. (2002). Eye movements and scene perception: Memory for things observed. *Perception & Psychophysics*, *64*(6), 882-895.

Isaacowitz, D. M., Toner, K., Goren, D., & Wilson, H.R. (2008) Looking while unhappy mood-congruent gaze in young adults, positive gaze in older adults. *Psychological Science*, *19*(9), 848–853. doi:10.1111/j.1467-9280.2008.02167.x

Isaacowitz, D. M., Wadlinger, H. A., Goren, D., & Wilson, H. R. (2006a). Is there an age-related positivity effect in visual attention? A comparison of two methodologies. *Emotion*, *6*(3), 511-516. doi:10.1037/1528-3542.6.3.511

Isaacowitz, D. M., Wadlinger, H. A., Goren, D., & Wilson, H. R. (2006b). Selective preference in visual fixation away from negative images in old age? An eye-tracking study. *Psychology and Aging*, *21*(1), 40-48. doi:10.1037/0882-7974.21.1.40

Itti, L. & Koch, C. (2000) A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10-12), 1489-1506.

Jackson, G., Aitken, C., & Roberts, P. (2015). *Guide 4. Case Assessment and Interpretation of Expert Evidence*. In *Communicating and Interpreting Statistical Evidence in the Administration of Criminal Justice*. Royal Statistical Society, United Kingdom. Retrieved from: <http://www.rss.org.uk/Images/PDF/influencing-change/rss-case-assessment-interpretation-expert-evidence.pdf>

Jacob, R.J.K. & Karn, K.S. (2003). Eye tracking in Human-Computer Interaction and usability research: Ready to deliver the promises. In: J. Hyönä, R. Radach and H. Deubel (Eds.) *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573-605) Amsterdam: Elsevier.

Jainta, S., Hoormann, J., & Jaschinski, W. (2008). Ocular accommodation and cognitive demand: An additional indicator besides pupil-size and cardiovascular measures? *Journal of Negative Results in BioMedicine*, 7(6). doi:10.1186/1477-5751-7-6

Jarodzka, H. (2010). *In the Eye of an Expert Conveying Perceptual Skills in Biological and Medical Domains via Eye Movement Modeling Examples*. (Unpublished Doctoral dissertation). University of Tübingen.

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction*, 20(2), 146-154.

Josephson, S. & Holmes, M.E. (2011) Selecting the suspect: An eye-tracking comparison of viewing of same-race vs. cross-race photographs in eyewitness identification. *Visual Communication Quarterly*, 18(4) 236-249.

Julian, R., Kely, S. & Robertson, J. (2012) "Get it right the first time": Critical issues at the crime scene. *Current Issues in Criminal Justice*, 24(1) 25-38.

Julian, R.D., Kely, S.F., Roux, C., Woodman, P., Robertson, J., Davey, A., Hayes, R., Margot, P., Ross, A., Sibly, H. & White, R. (2011) What is the value of forensic science? an overview of the effectiveness of forensic science in the Australian criminal justice system project. *Australian Journal of Forensic Sciences*, 43(4) 217-229. doi: 10.1080/00450618.2011.610820

Just, M.A., & Carpenter, P.A. (1987). *The psychology of reading and language comprehension*. Allyn & Bacon.

Just, M.A. & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4) 329-354.

Kamphuis, J. H., Emmelkamp, P. M., & Bartak, A. (2003). Individual differences in post-traumatic stress following post-intimate stalking: stalking severity and psychosocial variables. *British Journal of Clinical Psychology*, 42(2), 145-156.

Kaptein, R. G., & Van Gisbergen, J. A. (2006). Canal and otolith contributions to visual orientation constancy during sinusoidal roll rotation. *Journal of Neurophysiology*, 95(3), 1936-1948. doi:10.1152/jn.00856.2005

Kassin, S.M., Dror, I.E. & Kukucka, J. (2013) The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, 2(1) 42-52.

Kelty, S. F. (2011). Professionalism in Crime Scene Investigation: Recruitment strategies using the seven key attributes of top crime scene examiners. *Forensic Science Policy and Management: An International Journal*, 2(4), 198–204.

Kelty, S. F., Julian, R. & Robertson, J. (2011). Professionalism in crime scene investigation: The seven key attributes of top crime scene examiners. *Forensic Science Policy and Management: An International Journal*, 2(4), 175–186.

Kelty, S.F. & Gordon, H. (2012) Professionalism in Crime Scene Examination: Recruitment Strategies, Part 2: Using a Psychometric Profile of Top Crime Scene Examiners in Selection Decision Making. *Forensic Science Policy & Management: An International Journal*, 3(4), 189-199, doi: 10.1080/19409044.2013.858799

Kelty, S. F. & Gordon, H. (2015). No burnout at this coal-face: Managing occupational stress in forensic personnel and the implications for forensic and criminal justice agencies. *Psychiatry, Psychology and Law*, 22(2), 273-290. doi:10.1080/13218719.2014.941092

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2007). Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. *Journal of Memory and Language*, *56*(4), 575-591. doi:10.1016/j.jml.2006.05.004

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2006). Memory for specific visual details can be enhanced by negative arousing content. *Journal of Memory and Language*, *54*(1), 99-112.

Kensinger, E.A., & Schacter, D. L. (2006). When the Red Sox shocked the Yankees: Comparing negative and positive memories. *Psychonomic Bulletin and Review*, *13*, 757–763

Kerstholt, J.H., Paashuis, R. & Sjerps, M. (2007). Shoe print examinations: Effects of expectation, complexity and experience. *Forensic Science International*, *165*(1) 30-34.

Khan, R. S., Tien, G., Atkins, M. S., Zheng, B., Panton, O. N., & Meneghetti, A. T. (2012). Analysis of eye gaze: Do novice surgeons look at the same location as expert surgeons during a laparoscopic operation? *Surgical Endoscopy and Other Interventional Techniques*, *26*(12), 3536-3540. doi:10.1007/s00464-012-2400-7

Kimmel, D. L., Mammo, D., & Newsome, W. T. (2012). Tracking the eye non-invasively: simultaneous comparison of the scleral search coil and optical tracking techniques in the macaque monkey. *Frontiers of Behavioral Neuroscience*, *6*(49), 1-17. doi:10.3389/fnbeh.2012.00049

King, G. (2009). A framework of personal and environmental learning-based strategies to foster therapist expertise. *Learning in Health and Social Care*, 8(3), 185-199.

Kirby, C. E., Kennedy, Q., & Yang, J. H. (2014). Helicopter pilot scan techniques during low-altitude high-speed flight. *Aviation Space and Environmental Medicine*, 85(7), 740-744. doi:10.3357/ASEM.3888.2014

Kloosterman, A., Mapes, A., Geradts, Z., van Eijk, E., Koper, C., van den Berg, J., ... & van Asten, A. (2015). The interface between forensic science and technology: how technology could cause a paradigm shift in the role of forensic institutes in the criminal justice system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1674). doi:10.1098/rstb.2014.0264

Kocak, E., Ober, J., Berme, N., & Melvin, W. S. (2005). Eye motion parameters correlate with level of experience in video-assisted surgery: Objective testing of three tasks. *Journal of Laparoendoscopic and Advanced Surgical Techniques - Part A*, 15(6), 575-580. doi:10.1089/lap.2005.15.575

Koh, R.Y.I.; Park, T., Wickens, C. D.; Ong, L. T., & Chia, S.N. (2011) Differences in attentional strategies by novice and experienced operating theatre scrub nurses. *Journal of Experimental Psychology: Applied*, 17(3), 233-246. doi:10.1037/a0025171

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

Kuhn, G. & Findlay, J. M. (2010). Misdirection, attention and awareness: Inattention blindness reveals temporal relationship between eye movements and visual awareness. *Quarterly Journal of Experimental Psychology*, *63*(1), 136-146. doi:10.1080/17470210902846757

Kuhn, G., Tatler, B. W. & Cole, G. G. (2009). You look where I look! effect of gaze cues on overt and covert attention in misdirection. *Visual Cognition*, *17*(6-7), 925-944. doi:10.1080/13506280902826775

Kuhn, G., Tatler, B. W., Findlay, J. M., & Cole, G. G. (2008). Misdirection in magic: Implications for the relationship between eye gaze and attention. *Visual Cognition*, *16*(2-3), 391-405. doi:10.1080/13506280701479750

Kundel, H. L. & La Follette Jr, P. S. (1972). Visual search patterns and experience with radiological images. *Radiology*, *103*(3), 523–528.

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

Kurland, L., Gertner, A., Bartee, T., Chisholm, M., & McQuade, S. (2005) Using Cognitive Task Analysis and Eye Tracking to Understand Imagery Analysis. Retrieved from https://www.mitre.org/sites/default/files/pdf/05_1365.pdf

Lanata, A., Valenza, G., & Scilingo, E. P. (2013). Eye gaze patterns in emotional pictures. *Journal of Ambient Intelligence and Humanized Computing*, 4(6), 705-715.

Land, M. F., Mennie, N., & Rusted, J. (1999). Eye movements and the roles of vision in activities of daily living: making a cup of tea. *Perception*, 28(4), 1311-1328.

Land, M.F. (2006) Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, 25(3) 296-324.

Land, M. F. (2009). Vision, eye movements, and natural behavior. *Visual Neuroscience*, 26, 51–62. doi:10.1017/S0952523808080899.

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, 33(2) 363–374.

Lang, P. J., Bradley, M. M.; Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-8. University of Florida; Gainesville, FL.

Laubrock, J., Engbert, R., & Kliegl, R. (2005). Microsaccade dynamics during covert attention. *Vision Research*, 45(6), 721-730. doi:10.1016/j.visres.2004.09.029

Law, B., Atkins, M. S., Kirkpatrick, A. E., & Lomax, A. J. (2004). Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. *Eye Tracking Research and Applications Symposium (ETRA)*, 41-47.

Lawson, R. (2004). Depth rotation and mirror-image reflection reduce affective preference as well as recognition memory for pictures of novel objects. *Memory & Cognition*, 32(7), 1170-1181.

Lazarus, R. S. & Alfert, E. (1964). Short-circuiting of threat by experimentally altering cognitive appraisal. *The Journal of Abnormal and Social Psychology*, 69(2), 195-205. doi:10.1037/h0044635

Lee, H.C. & Pagliaro, E.M. (2013). Forensic Evidence and Crime Scene Investigation. *Journal of Forensic Investigation*. 1(1), 1-5.

Le Meur, O., & Baccino, T. (2013). Methods for comparing scanpaths and saliency maps: strengths and weaknesses. *Behavior Research Methods*, 45(1), 251-266. doi:10.3758/s13428-012-0226-9

Levi, D. (2008). Crowding - an essential bottleneck for object recognition: a mini-review. *Vision Research*, 48(5): 635-654. doi:10.1016/j.visres.2007.12.009.

Loftus, E. F., Loftus, G. R. & Messo, J. (1987). Some facts about “weapon focus”. *Law and Human Behavior*, 11(1), 55-62. doi:10.1007/BF01044839

Löw, A., Bradley, M. M., & Lang, P. J. (2013). Perceptual processing of natural scenes at rapid rates: effects of complexity, content, and emotional arousal. *Cognitive, Affective, & Behavioral Neuroscience*, 13(4), 860-868. doi:10.3758/s13415-013-0179-1

Lu, H. & Liu, Z. (2008). When a never-seen but less-occluded image is better recognized: Evidence from old–new memory experiments. *Journal of Vision*, 8(7), 1–9, doi:10.1167/8.7.31.

Lu, H. & Liu, Z. (2009). When a never-seen but less-occluded image is better recognized: Evidence from same-different matching experiments and a model. *Journal of Vision*, 9(4), 1–12, doi:10.1167/9.4.4.

Luce, R. D. (1986). *Response times: their role in inferring elementary mental organization*. New York: Oxford.

Ludwig, A., Fraser, J. & Williams, R. (2012) Crime Scene Examiners and Volume Crime Investigations: An Empirical Study of Perception and Practice. *Forensic Science Policy & Management: An International Journal*, 3(2), 53-61, doi: 10.1080/19409044.2012.728680

Lykken, D. T., & Venables, P. H. (1971). Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology*, 8(5), 656-672.)

Macdonald, R. G. & Tatler, B. W. (2013). Do as eye say: Gaze cueing and language in a real-world social interaction. *Journal of Vision*, 13(4), 6-6. doi:10.1167/13.4.6

MacKenzie, D. E., & Westwood, D. A. (2013). Observation patterns of dynamic occupational performance Modes d'observation du rendement occupationnel dynamique. *Canadian Journal of Occupational Therapy*, 80(2), 92-100.

Mann, D. Y., Williams, A. M., Ward, P., & Janelle, C. M. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport and Exercise Psychology*, 29(4), 457-478.

Mansour, J.K., Lindsay, R.C.L., Brewer, N. and Munhall, K.G. (2009) Characterizing visual behaviour in a lineup task. *Applied Cognitive Psychology*, 23(7) 1012-1026.

Marchewka, A., Żurawski, L., Jednoróg, K. & Grabowska, A. (2014) The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, 46(2), 596–610. doi:10.3758/s13428-013-0379-1

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Hubel, D. H. (2009). Microsaccades: a neurophysiological analysis. *Trends in Neurosciences*, 32(9), 463-475. doi: 10.1016/j.tins.2009.05.006

Mathot, S., Cristino, F., Gilchrist, I. D., & Theeuwes, J. (2012). A simple way to estimate similarity between pairs of eye movement sequences. *Journal of Eye Movement Research*, 5(1), 1–15.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314-324. doi:10.3758/s13428-011-0168-7

Mathôt, S., Siebold, A., Donk, M. & Vitu, F. (2015). Large pupils predict goal-driven eye movements. *Journal of Experimental Psychology: General*, 144(3), 513-521. doi: 10.1037/a0039168

McBride, M. F., & Burgman, M. A. (2012). What is expert knowledge, how is such knowledge gathered, and how do we use it to address questions in landscape ecology?. In A.H. Perera, C. A. Drew, C. J. Johnson (Eds.) *Expert knowledge and its application in landscape ecology* (pp. 11-38). Springer New York.

McCarley, J.S., Vais, M.J., Pringle, H., Kramer, A.F., Irwin, D.E. & Strayer, D.L. (2004) Conversation disrupts change detection in complex traffic scenes. *Human Factors*, 46(3) 424-436.

McConkie, G. W. (1997). Eye movement contingent display control: Personal reflections and comments. *Scientific Studies of Reading*, 1(4), 303–316.

McConkie, G.W. & Rayner, K. (1975) The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17(6), 578–586. doi:10.3758/BF03203972

McDonald, J.H. (2014). *Handbook of Biological Statistics* (3rd ed.). Sparky House Publishing.

McManus, I. C. (2013). iMAP and iMAP2 produce erroneous statistical maps of eye-movement differences. *Perception*, 42(10), 1075-1084.

Mearns, K., Flin, R., Gordon, R., & Fleming, M. (2001). Human and organizational factors in offshore safety. *Work & Stress*, 15(2), 144-160. doi: 10.1080/026783701102678370110066616

Megreya, A., Bindemann, M., Havard, C. & Burton, A. M. (2012) Identity-lineup location influences target selection: Evidence from eye movements. *Police and Criminal Psychology*, 27(2), 167–178.

Memmert, D., Simons, D.J. & Grimme, T. (2009). The relationship between visual attention and expertise in sports. *Psychology of Sport and Exercise*, 10(1), 146-151. doi:10.1016/j.psychsport.2008.06.002

Merlino, M.L., Freeman, T. M., Springer, V., Dahir, V., Hammond, D., Dyer, A. D., Found, B.J., Smith, L., & Duvall, I. (2014). Final report for the National Institute of Justice grant titled "Validity, Reliability, Accuracy, and Bias in Forensic Signature Identification." Submitted to National Institute of Justice, Washington, DC. Retrieved from <https://www.ncjrs.gov/pdffiles1/nij/grants/248565.pdf>

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

Miller, L. (2008). Criminal investigator stress: symptoms, syndromes, and practical coping strategies. *International Journal of Emergency Mental Health*, 11(2), 87-92.

Montagrin, A., Brosch, T., & Sander, D. (2013). Goal conduciveness as a key determinant of memory facilitation. *Emotion*, 13(4), 622-628. doi:10.1037/a0033066

Moreno, F., Reina, R., Luis, V., & Sabido, R. (2002). Visual search strategies in experienced and inexperienced gymnastic coaches. *Perceptual and Motor skills*, 95(3), 901–902.

Morimoto, C. H., & Mimica, M. R. (2005). Eye gaze tracking techniques for interactive applications. *Computer Vision and Image Understanding*, 98(1), 4-24. doi:10.1016/j.cviu.2004.07.010

Morphew, J. W., Mestre, J. P., Ross, B. H. & Strand, N. E. (2015). Do experts and novices direct attention differently in examining physics diagrams? A study of change detection using the flicker technique. *Physical Review Special Topics - Physics Education Research*, 11(2) doi:10.1103/PhysRevSTPER.11.020104

Morrison, B. W., Wiggins, M. W., & Porter, G. (2010). User preference for a control-based reduced processing decision support interface. *International Journal of Human-Computer Interaction*, 26(4), 297-316. doi:10.1080/10447310903575465

Munoz, D. P., & Everling, S. (2004). Look away: the anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, 5(3), 218-228.

Murphy, P. J., Duncan, A. L., Glennie, A. J., & Knox, P. C. (2001). The effect of scleral search coil lens wear on the eye. *British Journal of Ophthalmology*, 85(3), 332-335. doi:10.1136/bjo.85.3.332

Nakano, T., Kato, M., Morito, Y., Itoi, S., & Kitazawa, S. (2013). Blink-related momentary activation of the default mode network while viewing videos. *Proceedings of the National Academy of Sciences of the United States of America*, 110(2), 702-706. doi:10.1073/pnas.1214804110

Nakhaeizadeh, S., Dror, I.E. & Morgan, R.M. (2014) Cognitive bias in forensic anthropology: Visual assessment of skeletal remains is susceptible to confirmation bias. *Science and Justice*, 54(3) 208-214.

National Research Council of the National Academy of Sciences (2009). *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: National Academy of Sciences.

Nelson, K., Laney, C., Fowler, N.B., Knowles, E.D., Davis, D. & Loftus, E. (2011). Change blindness can cause mistaken eyewitness identification. *Legal & Criminological Psychology*, 16(1), 62-74. doi: 10.1348/135532509X482625

Newell, A. & Simon, H. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall.

Nicolaou, M., James, A., Darzi, A., & Yang, G. Z. (2004). A Study of saccade transition for attention segregation and task strategy in laparoscopic surgery. In *Medical image computing and computer-assisted intervention—MICCAI 2004* (pp. 97-104). Springer Berlin Heidelberg.

Nodine, C. F. & Kundel, H. L. (1987). Using eye movements to study visual search and to improve tumor detection. *Radiographics*, 7(6), 1241–1250.

Nodine, C. F., Kundel, H. L., Lauver, S. C., & Toto, L. C. (1996). Nature of expertise in searching mammograms for breast masses. *Academic Radiology*, 3(12), 1000–1006. doi:10.1016/S1076-6332(96)80032-8

Norman, G. R., Coblenz, C. L., Brooks, L. R., & Babcock, C. J. (1992). Expertise in visual diagnosis: a review of the literature. *Academic Medicine*, 67(10), S78-83.

Nuthmann, A. & van der Meer, E.E. (2005). Time's arrow and pupillary response. *Psychophysiology*, 42(3), 306-317. doi:10.1111/j.1469-8986.2005.00291.x

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: a computational model of fixation durations in scene viewing. *Psychological Review*, 117(2), 382-405. doi:10.1037/a0018924

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*, 42(1), 188-204. doi:10.3758/BRM.42.1.188

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, 14(8), 1215-1229. doi:10.1162/089892902760807212

Olf, M., Langeland, W., Draijer, N., & Gersons, B. P. R. (2007). Gender differences in posttraumatic stress disorder. *Psychological Bulletin*, 133(2), 183-204. doi:10.1037/0033-2909.133.2.183

O'Meara, P. , Munro, G. , Cooper, S., Williams, B., Bogossian, F., Ross, L., Sparkes, L. , Browning, M., McClounan, M. (2015). Developing situation awareness amongst nursing and paramedicine students utilizing eye tracking technology and video debriefing techniques: A proof of concept paper. *International Emergency Nursing*, 23(2), 94-99. doi:10.1016/j.ienj.2014.11.001

Omodei, M. M. & McLennan, J. (1994). Studying complex decision making in natural settings: Using a head-mounted video camera to study competitive orienteering. *Perceptual and Motor Skills*, 79(3 Pt 2) 1411-1425.

Orquin, J. L. & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, 144(1) 190-206.

Overney, L. S., Blanke, O., & Herzog, M. H. (2008). Enhanced temporal but not attentional processing in expert tennis players. *Plos One*, 3(6) doi:10.1371/journal.pone.0002380.

Parris, B. A., Kuhn, G., Mizon, G. A., Benattayallah, A., & Hodgson, T. L. (2009). Imaging the impossible: An fMRI study of impossible causal relationships in magic tricks. *Neuroimage*, 45(3), 1033–1039. doi:10.1016/j.neuroimage.2008.12.036

Partala, T. & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human Computer Studies*, 59(1-2), 185-198. doi:10.1016/S1071-5819(03)00017-X

Patel, V.L , Arocha, .J.F. & Zhang, J. (2012). Medical Reasoning and Thinking. In K. J. Holyoak & R. G. Morrison, (Eds.) *The Oxford Handbook of Thinking and Reasoning* (pp.736-754). Oxford University Press, New York

Phelps, E. A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annual Review of Psychology*, 57, 27–53.

Pickel, K. L. (1998). Unusualness and threat as possible causes of 'weapon focus'. *Memory*, 6, 277–295

Piras, A., Lobietti, R., & Squatrito, S. (2014). Response time visual search strategy, and anticipatory skills in volleyball players. *Journal of Ophthalmology*, 2014,1–10. doi: 10.1155/2014/189268

Piras, A., Pierantozzi, E., & Squatrito, S. (2014). Visual search strategy in judo fighters during the execution of the first grip. *International Journal of Sports Science and Coaching*, 9(1), 185-197. doi:10.1260/1747-9541.9.1.185

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, 343(6255), 263–266

PoliceOne. (2015, August 24). *How 3D laser scanners are changing crime scene investigations*. Retrieved from <https://www.policeone.com/police-products/3D-Laser-Scanners/articles/8720378-How-3D-laser-scanners-are-changing-crime-scene-investigations/>

Positive Science (2015). *Eye Tracking Software User Guides and Hardware Documentation*. New York, NY: Positive Science, LLC.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3-25.

Rayner, K. (1975) The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7(1), 65–81. doi:10.1016/0010-0285(75)90005-5

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin*, 85(3), 618-660. doi:10.1037/0033-2909.85.3.618

Rayner, K. (1998) Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3) 372-422. doi:10.1037/0033-2909.124.3.372

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of reading (2nd ed.)*. Psychology Press.

Reader, T., Flin, R., Lauche, K., & Cuthbertson, B. H. (2006). Non-technical skills in the intensive care unit. *British Journal of Anaesthesia*, 96(5), 551-559. doi:10.1093/bja/ae1067

Reder, S.M. (1973) On-line monitoring of eye position signals in contingent and noncontingent paradigms. *Behavior Research Methods & Instrumentation*, 5(2), 218-228. doi:10.3758/BF03200168

Reed, A. E., Chan, L., & Mikels, J. A. (2014). Meta-analysis of the age-related positivity effect: Age differences in preferences for positive over negative information. *Psychology and Aging*, 29(1), 1-15. doi:10.1037/a0035194

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, 26(4), 445-476. doi:10.1017/S0140525X03000104

Reingold, E.M., Charness, N., Pomplun, M. & Stampe, D.M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, 12(1) 48-55.

Reingold E.M., & Sheridan H (2011). Eye movements and visual expertise in chess and medicine. In S.P. Liversedge, I.D. Gilchrist, S. Everling, (Eds.). *Oxford handbook on eye movements* (pp.767–786). Oxford, UK: Oxford University Press.

Rensink, R.A. (2008). On the applications of change blindness. *Psychologia*, 51(2) 100-106.

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(5) 368-373.

Ribaux, O., Baylon, A., Roux, C., Delémont, O., Lock, E., Zingg, C. & Margot, P. (2010a). Intelligence-led crime scene processing. part I: Forensic intelligence. *Forensic Science International*, 195(1-3) 10-16.

Ribaux, O., Baylon, A., Lock, E., Delémont, O., Roux, C., Zingg, C. & Margot, P. (2010b). Intelligence-led crime scene processing. part II: Intelligence and crime scene examination. *Forensic Science International*, 199(1-3) 63-71.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29(6), 1045-1060. doi:10.1207/s15516709cog0000_29

Richardson, M. J., Lopresti-Goodman, S., Mancini, M., Kay, B. A., & Schmidt, R. C. (2008). Comparing the attractor strength of intra- and interpersonal interlimb coordination using cross recurrence analysis. *Neuroscience Letters*, 438(3), 340-345. doi:10.1016/j.neulet.2008.04.083.

Richardson, M. J., Schmidt, R. C., & Kay, B. A. (2007). Distinguishing the noise and attractor strength of coordinated limb movements using recurrence analysis. *Biological Cybernetics*, 96(1), 59-78. doi:10.1007/s00422-006-0104-6.

Richardson, D. C., Street, C. N. H., Tan, J. Y. M., Kirkham, N. Z., Hoover, M. A., & Ghane Cavanaugh, A. (2012). Joint perception: gaze and social context. *Frontiers in Human Neuroscience*, 6, 194. doi:10.3389/fnhum.2012.00194

Richstone, L., Schwartz, M. J., Seideman, C., Cadeddu, J., Marshall, S., & Kavoussi, L. R. (2010). Eye metrics as an objective assessment of surgical skill. *Annals of Surgery*, 252(1), 177-182. doi:10.1097/SLA.0b013e3181e464fb

Riggs, L., McQuiggan, D. A., Anderson, A. K., & Ryan, J. D. (2010). Eye movement monitoring reveals differential influences of emotion on memory. *Frontiers in Psychology*, 1(NOV). doi:10.3389/fpsyg.2010.00205

Roberts, P. & Aitken, C. (2014). *Guide 3. The Logic of Forensic Proof: Inferential Reasoning in Criminal Evidence and Forensic Science*. Royal Statistical Society, United Kingdom. Retrieved from <http://www.maths.ed.ac.uk/~cgga/Guide-3-WEB.pdf>

Robertson, D.J., Noyes, E., Dowsett, A.J., Jenkins, R., & Burton, A.M. (2016). Face Recognition by Metropolitan Police Super-Recognisers. *PLoS One* 11(2): e0150036. doi:10.1371/journal.pone.0150036

Robinson, D. A. (1965). The mechanics of human smooth pursuit eye movement. *The Journal of Physiology*, 180(3), 569-591.

Robinson, D. A., Gordon, J. L., & Gordon, S. E. (1986). A model of the smooth pursuit eye movement system. *Biological Cybernetics*, 55(1), 43-57. doi:10.1007/BF00363977

Rolfs, M. (2009). Microsaccades: small steps on a long way. *Vision Research*, 49(20), 2415-2441. doi:10.1016/j.visres.2009.08.010

Rosengrant, D., Thomson, C. & Mzoughi, T. (2009) Comparing Experts and Novices in Solving Electrical Circuit Problems with the Help of Eye-Tracking. *AIP Conference Proceedings*, 1179(1), 249-252. doi:10.1063/1.3266728

Salvucci, D. & Goldberg, J. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the Eye Tracking Research and Applications Symposium*, New York, 71–78.

Santella, A., & DeCarlo, D. (2004). Robust Clustering of Eye Movement Recordings for Quantification of Visual Interest. *Eye Tracking Research and Applications Symposium (ETRA)*, San Antonio, TX, 27–34

Santtila, P., Korpela, S., & Hakkanen, H. (2004). Expertise and decision-making in the linking of car crime series. *Psychology, Crime and Law*, 10(2), 97–112. doi:10.1080/1068316021000030559

Savelsbergh, G. J. P., Williams, A. M., van der Kamp, J., & Ward, P. (2002). Visual search, anticipation and expertise in soccer goalkeepers. *Journal of Sports Sciences*, 20(3), 279–287. doi:10.1080/026404102317284826

Schaible, L. M. & Gecas, V. (2010). The impact of emotional labor and value dissonance on burnout among police officers. *Police Quarterly*, 13(3), 316-341. doi: 10.1177/1098611110373997

Schienze, A., Übel, S., Gremsl, A., Schöngassner, F., & Körner, C. (2016). Disgust Proneness and the Perception of Disgust-Evoking Pictures: An Eye-Tracking Experiment. *Journal of Psychophysiology*. doi:10.1027/0269-8803/a000162

Schraagen, J. M. (2006). Task analysis. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. Hoffman (Eds.). *The Cambridge Handbook of Expertise and Expert Performance* (pp. 185-201). New York: Cambridge University Press.

Schraagen, J. M. & Lejjenhorst, H. (2001). Searching for evidence: knowledge and search strategies used by forensic scientists. In: E. Salas and G. Klein (ed.) *Linking Expertise and Naturalistic Decision Making* (pp.263-274). Mahwah, NJ: Lawrence Erlbaum Associates.

Schiffer, B. (2009). *The Relationship between Forensic Science and Judicial Error: a Study Covering Error Sources, Bias, and Remedies* (Unpublished Doctoral dissertation). University of Lausanne.

Schrivers, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors*, 50(6), 864–878. doi:10.1518/001872008X374974

Schuliar, Y. & Crispino, F. (2015) Semiotics, Heuristics and Inferences Used by Forensic Scientists. In M.M. Houck (Ed.) *Professional Issues in Forensic Sciences* (pp. 105-109). Oxford: Academic Press.

Sewell, J. D. (1994). The stress of homicide investigations. *Death studies*,18(6), 565-582. doi:10.1080/07481189408252701

Shank, M. D., & Haywood, K. M. (1987). Eye movements while viewing a baseball pitch. *Perceptual and Motor Skills*, 64(3c), 1191–1197.

Sherman, L. (2013). The Rise of Evidence-Based Policing: Targeting, Testing, and Tracking. In M. Tonry (Ed.) ‘*Crime and Justice in America, 1975-2025*’, *Crime and Justice*, 42, (pp.377-451). Chicago, University of Chicago Press.

- Simons, D.J. (2000). Current approaches to change blindness. *Visual Cognition*, 7(1-3) 1-15.
- Simons, D. J., Franconeri, S. L., & Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception*, 29(10), 1143–1154.
- Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in cognitive sciences*, 1(7), 261–267.
- Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a realworld interaction. *Psychonomic Bulletin & Review*, 5(4), 644-649.
- Sita, J., Found, B. & Rogers, D.K. (2002). Forensic handwriting examiners' expertise for signature comparison. *Journal of Forensic Sciences*, 47(5) 1117-1124.
- Smart, S. M., Berry, M. A., & Rodriguez, D. N. (2014). Skilled observation and change blindness: A comparison of law enforcement and student samples. *Applied Cognitive Psychology*, 28(4), 590-596. doi:10.1002/acp.3021
- Smith, L. & Bond, J. (2014) *Criminal Justice and Forensic Science: A Multidisciplinary Introduction*. Palgrave Macmillan.
- Smith, E. L., Grabowecky, M., & Suzuki, S. (2008). Self awareness affects vision. *Current Biology*, 18(10), R414–R415. doi:10.1016/j.cub.2008.03.009.

Snyder, P. J., Nussbaum, P. D., & Robins, D. L. (2006). *Clinical neuropsychology: A pocket handbook for assessment*. American Psychological Association.

Sodergren, M. H., Orihuela-Espina, F., Clark, J., Teare, J., Yang, G. Z., & Darzi, A. (2010). Evaluation of orientation strategies in laparoscopic cholecystectomy. *Annals of Surgery*, 252(6), 1027–1036. doi:10.1097/SLA.0b013e3181e49683

Song, Y., Dixon, S. Pearce, M. T. & Halpern, A.R. (2016). Perceived and induced emotion responses to popular music: Categorical and dimensional models. *Music Perception: An Interdisciplinary Journal*, 33(4), 472-492.

Spivey, M. J., & Dale, R. (2011). Eye movements both reveal and influence problem solving. In S. P. Liversedge, I. D. Gilchrist, & S. Everding (Eds.). *The Oxford handbook of eye movements* (pp. 551-562). Oxford, UK: Oxford University Press.

SR Research (2010). *Eyelink 1000 User's Manual*, Version 1.5.2. Mississauga, ON: SR Research Ltd.

Stein, M. B., Jang, K. L., Taylor, S., Vernon, P. A., & Livesley, W. J. (2002). Genetic and environmental influences on trauma exposure and posttraumatic stress disorder symptoms: a twin study. *American Journal of Psychiatry*, 159(10), 1675-1681. doi: 10.1176/appi.ajp.159.10.1675

Stewart, N., Hermens, F., & Matthews, W.J. (2016). Eye movements in risky choice. *Journal of Behavioral Decision Making*, 29(2-3), 116-136. doi: 10.1002/bdm.1854.

Swets, J. A. (2014). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. Psychology Press.

Tangen, J. M. (2013) Identification personified. *Australian Journal of Forensic Sciences*, 45 (3), 315-322. doi:10.1080/00450618.2013.782339

Tatler, B. W. (2014) Eye Movements from Laboratory to Life. In: M. Horsley, M. Elliot, M. A. Knight & R. Reilly (eds.) *Current Trends in Eye Tracking Research* (pp.17-36). Switzerland: Springer International Publishing.

Tatler, B. W., Hayhoe, M. M., Land, M. F., Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision*, 11(5), 1–23, doi:10.1167/11.5.5.

Tatler, B. W., & Vincent, B. (2008). Systematic tendencies in scene viewing. *Journal of Eye Movement Research*, 2(2):5, 1-18.

The University of Edinburgh (2009). *The DIEM Project*. Retrieved from <http://www.psy.ed.ac.uk/diem/index.html>

Tien, T., Pucher, P. H., Sodergren, M. H., Sriskandarajah, K., Yang, G. Z., & Darzi, A. (2014). Eye tracking for skills assessment and training: a systematic review. *Journal of Surgical Research*, 191(1), 169-178.

Tien, G., Atkins, M. S., & Zheng, B. (2012). Measuring gaze overlap on videos between multiple observers. *Eye Tracking Research and Applications Symposium (ETRA)*, 309-312. doi:10.1145/2168556.2168623

Tien, G., Atkins, M. S., Zheng, B., & Swindells, C. (2010). Measuring situation awareness of surgeons in laparoscopic training. *Eye Tracking Research and Applications Symposium (ETRA)*, 149-152. doi:10.1145/1743666.1743703

Tien, T., Pucher, P. H., Sodergren, M. H., Sriskandarajah, K., Yang, G., & Darzi, A. (2014). Eye tracking for skills assessment and training: A systematic review. *Journal of Surgical Research*, 191(1), 169-178. doi:10.1016/j.jss.2014.04.032

Townsend, J. T., & Ashby, F. G. (1983). *The Stochastic Modeling of Elementary Psychological Processes*. Cambridge: Cambridge University Press

Trappenberg, T. P., Dorris, M. C., Munoz, D. P., & Klein, R. M. (2001). A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. *Journal of Cognitive Neuroscience*, 13(2), 256-271. doi:10.1162/089892901564306

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 353(1373), 1295-1306. doi:10.1098/rstb.1998.0284

Trickett, S., & Traflet, J. (2007). A primer on verbal protocol analysis. In D. Schmorow, J. Cohn and D. Nicholson (Eds.), *The PSI Handbook of Virtual Environments for Training and Education: Developments for the Military and Beyond* (pp.332-346). Westport, CT Praeger Security International

Tseng, P. H., Carmi, R., Cameron, I. G. M., Munoz, D. P. & Itti, L. (2009) Quantifying center bias of observers in free viewing of dynamic natural scenes. *Journal of Vision*, 9(7), 1–16, doi:10.1167/9.7.4.

Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, 12 (3), 473–494.

Vaidyanathan, P., Pelz, J., Alm, C., Pengcheng, S., & Haake, A. (2014) Recurrence Quantification Analysis Reveals Eye-movement Behavior Differences Between Experts and Novices. *Eye Tracking Research and Applications Symposium (ETRA)*, 303-306.

Van den Eeden, C. A. J., de Koop, C. J., & van Koppen, P. J. (In Press) Forensic expectations: Investigating a crime scenewith prior information, *Science and Justice*, doi:10.1016/j.scijus.2016.08.003

Van der Geest, J. N., & Frens, M. A. (2002). Recording eye movements with video-oculography and scleral search coils: a direct comparison of two methods. *Journal of Neuroscience Methods*, 114(2), 185-195. doi:10.1016/S0165-0270(01)00527-1

Van Gog, T., Kester, L., Nievelein, F., Giesbers, B., & Paas, F. (2009). Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. *Computers in Human Behavior*, 25, 325-331.

Van Gog, T., Paas, F., van Merriënboer, J.J. & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, 11(4), 237-244.

Van Meeuwen, L. W., Jarodzka, H., Brand-Gruwel, S., Kirschner, P. A., de Bock, J. J. & van Merriënboer, J. J. (2014). Identification of effective visual problem solving strategies in a complex visual domain. *Learning and Instruction*, 32, 10-21.
doi:10.1016/j.learninstruc.2014.01.004

Van Paasschen, J., Bacci, F., & Melcher, D.P. (2015) The Influence of Art Expertise and Training on Emotion and Preference Ratings for Representational and Abstract Artworks. *PLoS One* 10(8): e0134241. doi:10.1371/journal.pone.0134241

Vickers, J. N. (1996). Visual control when aiming at a far target. *Journal of Experimental Psychology. Human Perception and Performance*, 22(2), 342–354

Vickers, J. N. (2016). Origins and current issues in Quiet Eye research. *Current Issues in Sport Science*, 1,101. doi: 10.15203/CISS_2016.101. *Learning and Instruction*, 32, 10-21.
doi:10.1016/j.learninstruc.2014.01.004

Vine, S. J., Chaytor, R. J., McGrath, J. S., Masters, R. S., & Wilson, M. R. (2013). Gaze training improves the retention and transfer of laparoscopic technical skills in novices. *Surgical Endoscopy and Other Interventional Techniques*, 27(9), 3205-3213. doi:10.1007/s00464-013-2893-8

Vivona, B. D. (2014). Humor functions within crime scene investigations: group dynamics, stress, and the negotiation of emotions. *Police Quarterly*, 17(2), 127-149. doi:10.1177/1098611114531418

Von Hofsten, C., & Rosander, K. (1997). Development of smooth pursuit tracking in young infants. *Vision Research*, 37(13), 1799-1810. doi:10.1016/S0042-6989(96)00332-X

Wade N. J., & Tatler B. W. (2005). *The Moving Tablet of the Eye*. Oxford, UK: Oxford University Press.

Wang J. T. (2011). Pupil dilation and eye-tracking, In M. Schulte-Mecklenbeck, A. Kuehberger & R. Ranyard (Eds.) *A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide*, (pp.185-194). New York, NY: Psychology Press.

Werner, S. & Thies, B. (2000). Is 'change blindness' attenuated by domain-specific expertise? An expert-novice's comparison of change detection in football images. *Visual Cognition*, 7(1-3) 163-173

Whitney, D., & Levi, D. M. (2011). Visual Crowding: a fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, 15(4), 160–168. <http://doi.org/10.1016/j.tics.2011.02.005>

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance (3rd ed.)*. Upper Saddle River, NJ: Prentice Hall

Wickens, C., Gordon, S. & Liu, Y. (1998). *An Introduction to Human Factors Engineering*. United States: Addison-Wesley Educational Publishers Inc.

Williams, A.M. & Davids, K. (1998). Visual search strategy, selective attention, and expertise in soccer. *Research Quarterly for Exercise and Sport*, 69(2), 111-128.

Williams, A.M., Davids, K., Burwitz, L., & Williams, J.G. (1994). Visual search strategies in experienced and inexperienced soccer players. *Research Quarterly for Exercise and Sport*, 65(2), 127-135. doi:10.1080/02701367.1994.10607607

Williams, R. & Weetman, J. (2013) Enacting forensics in homicide investigations. *Policing and Society*, 23(3), 376-389. doi:10.1080/10439463.2012.703200

Wilson, M., Coleman, M., & McGrath, J. (2010). Developing basic hand-eye coordination skills for laparoscopic surgery using gaze training. *BJU International*, 105(10), 1356–1358. doi:10.1111/j.1464-410X.2010.09315.x

- Wilson, M., McGrath, J., Vine, S., Brewer, J., Defriend, D. & Masters, R. (2010). Psychomotor control in a virtual laparoscopic surgery training environment: gaze control parameters differentiate novices from experts. *Surgical Endoscopy and Other Interventional Techniques*, 24(10), 2458-2464. doi:10.1007/s00464-010-0986-1
- Wilson, M. R., McGrath, J. S., Vine, S. J., Brewer, J., Defriend, D., & Masters, R. S. (2011). Perceptual impairment and psychomotor control in virtual laparoscopic surgery. *Surgical Endoscopy and Other Interventional Techniques*, 25(7), 2268-2274. doi:10.1007/s00464-010-1546-4
- Wilson-Kovacs, D. (2013). 'Backroom boys': Occupational dynamics in crime scene examination. *Sociology*, 48(4) 763-779.
- Winkielman, P., Knutson, B., Paulus, M., & Trujillo, J. L. (2007). Affective influence on judgments and decisions: Moving towards core mechanisms. *Review of General Psychology*, 11(2), 179-192. doi:10.1037/1089-2680.11.2.179.
- Wolfe, J.M. (1994). Guided search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, 1(2), 202-238.
- Wood, G., & Wilson, M. R. (2011). Quiet-eye training for soccer penalty kicks. *Cognitive Processing*, 12(3), 257-266. doi:10.1007/s10339-011-0393-0.

Wu, D.W.L., Anderson, N. C., Bischof, W. F. & Kingstone, A. (2014). Temporal dynamics of eye movements are related to differences in scene complexity and clutter. *Journal of Vision, 14*(9), 1–14. doi:10.1167/14.9.8

Yarbus, A L, (1967). *Eye Movements and Vision*. New York: Plenum Press.

Wyatt, D. (2014). Practising crime scene investigation: trace and contamination in routine work, *Policing and Society: An International Journal of Research and Policy, 24*(4), 443-458, doi: 10.1080/10439463.2013.868460

Yu, C., Wang, E. M., Li, W., Braithwaite, G., & Greaves, M. (2016). Pilots' visual scan patterns and attention distribution during the pursuit of a dynamic target. *Aerospace Medicine and Human Performance, 87*(1), 40-47. doi:10.3357/AMHP.4209.2016

Yule, S., Flin, R., Maran, N., Rowley, D. R., Youngson, G., & Paterson-Brown, S. (2008). Surgeons' non-technical skills in the operating room: Reliability testing of the NOTSS behaviour rating system. *World Journal of Surgery, 32*(4), 548-556. doi:10.1007/s00268-007-9320-z

Yule, S., Flin, R., Paterson-Brown, S., & Maran, N. (2006). Development of a rating system for surgeons' non-technical skills. *Medical Education, 40*(11), 1098-1104. doi:10.1111/j.1365-2929.2006.02610.x

Zelinsky, G., & Sheinberg, D. (1995). Why some search tasks take longer than others: Using eye movements to redefine reaction times. *Studies in Visual Information Processing*, 6, 325-336.

Zheng, B., Jiang, X., Tien, G., Meneghetti, A., Panton, O.N., & Atkins, M.S. (2012) Workload assessment of surgeons: correlation between NASA TLX and blinks. *Surgical Endoscopy and Other Interventional Techniques*, 26(10), 2746-2750. doi:10.1007/s00464-012-2268-6

Appendix A

Information Letter Sent to the Managers about the study

Perception and Decision Making processes in Crime Scene Investigators (CSIs) Using Eye-tracking devices

Dear Sir/Lady,

I am writing to request your consideration in regard of participation in a research project being conducted by the University of Lincoln. The research is being conducted as part of a programme of work focusing on Crime Scene Examination and the decision-making of professionals tasked to assess scenes (Crime Scene Investigators – CSIs). The research team comprises Dr Jose Gonzalez-Rodriguez and Professor Timothy Hodgson and me, Murat Ozger. The research programme is "Attention to evidence: A comparison of perception and decision making processes in novice and experienced Crime Scene Investigators (CSIs) using eye tracking devices in simulated crime scene scenarios."

In order to meet the inclusion criteria for this research, we are looking to recruit novice and experienced CSIs to participate in this research project. I am asking for your permission to access CSIs in your department and ask about volunteers to participate for this project.

The project proposes to examine whether there are differences in visual perception, attention, search strategies and decision-making processes in a group of novice and experienced CSIs regarding the evidences, environment and consideration of perpetrators' activities through examining different simulated and manipulated crime scene scenarios (esp. serious crime scene scenarios, but also volume crime scenarios), which will also contribute to the understanding of the contextual effects of the scene on CSIs.

It is hoped that the findings of this research will inform tailored training initiatives to accelerate the novice CSIs induction time after finishing the trainings for volume and serious crime scene examinations by modeling optimum search strategies, competence and decision making processes to perform as effectively as the experienced CSIs. In addition to the present training programs, it is hoped to set up a new training tool based on the data from the experiments.

All participant information, including their originating police forces will remain anonymous and confidential and all data will be stored in a secure environment, although will not be identifiable to any individuals. Only the researcher will have access to the raw data. All sessions of the study will be held in special labs at the University of Lincoln.

Should you have any particular questions about the nature and content of this research study, please contact the lead researcher Murat Ozger (see contact details below). It will also be possible to arrange informal information sessions in your office/station to answer queries and to explain the project further, if you would like to discuss the research individually.

Your permission is greatly valued.

If you have any questions please contact me on **01522 837366** or my mobile **07465 062252** or e-mail 12378768@students.lincoln.ac.uk

The contact details of the research team are :

Dr Jose Gonzalez Rodriguez
School of Life Sciences
Phone :**01522 886878**
E-mail : jgonzalezrodriguez@lincoln.ac.uk

Prof Timothy Hodgson
School of Psychology
Phone :**01522 886159**
E-mail : tlhodgson@lincoln.ac.uk

Yours sincerely,

Murat OZGER

Doctoral Student
University of Lincoln
School of Life Sciences / School of Psychology
Brayford Pool
Lincoln
LN6 7TS

This research study has been approved by the University of Lincoln Research Ethics Committee. You may contact the Ethics Committee through the Senior Administrator, Mrs Judy Steven, phone **01522 83 5511, e-mail : js Steven@lincoln.ac.uk**

Appendix B

List of Details of the Stimuli used in Chapter 5

No	Image Type	ID	Description		Change	Location of the change in Image
1	Crime Scene	CS1	Room	Adaptor	Delete	Top-right
2	Crime Scene	CS2	Room	Drawer	Delete	Bottom-left
3	Crime Scene	CS3	Kitchen	Sugar Jar	Delete	Top-left
4	Crime Scene	CS4	Parking Lot	Parked car	Add	Top-right
5	Crime Scene	CS5	Room	Mirror-reflection	Delete	Top-right
6	Crime Scene	CS6	Room	Knife	Change Location	Bottom-right
7	Crime Scene	CS7	Room	Can	Delete	Bottom-right
8	Crime Scene	CS8	Building	Window	Delete	Top-right
9	Crime Scene	CS9	Room	Cable	Delete	Top-left
10	Crime Scene	CS10	Kitchen	Shadow	Delete	Top-left
11	Crime Scene	CS11	Garden	Cloth	Delete	Top-left
12	Crime Scene	CS12	Room	Shoemarks	Add	Bottom-right
13	Crime Scene	CS13	Toilet	Shampoo	Change Location	Bottom-left
14	Crime Scene	CS14	Room	Glasses Case	Change Location	Top-right
15	Natural Scene	NAT1	Traffic Scene	Traffic Light	Change Location	Bottom-left
16	Natural Scene	NAT2	Parking Lot	Manhole	Add	Bottom-right
17	Natural Scene	NAT3	Building	Cathedral	Delete	Top-left
18	Natural Scene	NAT4	Train Station	Train Window	Add	Bottom-left
19	Natural Scene	NAT5	Traffic Scene	Person	Change Location	Bottom-right
20	Natural Scene	NAT6	Room	Person	Delete	Top-left
21	Natural Scene	NAT7	Traffic Scene	Sitting People	Delete	Bottom-right
22	Natural Scene	NAT8	Traffic Scene	Traffic Sign	Delete	Bottom-left
23	Natural Scene	NAT9	Parking Lot	Person	Change Location	Bottom-left
24	Natural Scene	NAT10	Building	Flag	Add	Top-right
25	Natural Scene	NAT11	Building	Castle	Delete	Top-left
26	Natural Scene	NAT12	Traffic Scene	Traffic Sign	Change Location	Top-right
27	Natural Scene	NAT13	Traffic Scene	Bin	Delete	Bottom-left
28	Natural Scene	NAT14	Building	Office Sign	Delete	Top-left

Appendix C

List of Details of the Stimuli used in Chapter 7

No	Database	ID	Category	Emotional Content	Vertical/Horizontal	Description		Luminance
1	SURVEY	CN1	Crime Scene	NEGATIVE	v	Corridor	INDOOR	99.90
2	SURVEY	CN2	Crime Scene	NEGATIVE	h	Room	INDOOR	99.53
3	SURVEY	CN3	Crime Scene	NEGATIVE	h	Kitchen	INDOOR	88.71
4	SURVEY	CN4	Crime Scene	NEGATIVE	h	Room	INDOOR	99.00
5	SURVEY	CN5	Crime Scene	NEGATIVE	h	Sink	INDOOR	134.49
6	SURVEY	CN6	Crime Scene	NEGATIVE	h	Gun	INDOOR	107.90
7	SURVEY	CN7	Crime Scene	NEGATIVE	h	Room	INDOOR	97.66
8	SURVEY	CN8	Crime Scene	NEGATIVE	h	Room	INDOOR	105.53
9	SURVEY	CN9	Crime Scene	NEGATIVE	h	Bedroom	INDOOR	81.11
10	SURVEY	CN10	Crime Scene	NEGATIVE	h	Toilet	INDOOR	94.92
11	SURVEY	CN11	Crime Scene	NEGATIVE	h	Gun	OUTDOOR	73.28
12	SURVEY	CN12	Crime Scene	NEGATIVE	h	Room	INDOOR	101.00
13	SURVEY	CN13	Crime Scene	NEGATIVE	h	Room	INDOOR	109.81
14	SURVEY	CN14	Crime Scene	NEGATIVE	h	Bedroom	INDOOR	133.42
15	SURVEY	CN15	Crime Scene	NEGATIVE	h	Kitchen	INDOOR	119.40
16	SURVEY	CS1	Crime Scene	NEUTRAL	v	Bedroom	INDOOR	121.59
17	SURVEY	CS2	Crime Scene	NEUTRAL	h	Bedroom	INDOOR	96.97
18	SURVEY	CS3	Crime Scene	NEUTRAL	h	Bedroom	INDOOR	82.89
19	SURVEY	CS4	Crime Scene	NEUTRAL	h	Room	INDOOR	119.40
20	SURVEY	CS5	Crime Scene	NEUTRAL	h	Room	INDOOR	96.87
21	SURVEY	CS6	Crime Scene	NEUTRAL	h	Kitchen	INDOOR	109.01
22	SURVEY	CS7	Crime Scene	NEUTRAL	h	Table	INDOOR	76.51
23	SURVEY	CS8	Crime Scene	NEUTRAL	h	Room	INDOOR	103.97
24	SURVEY	CS9	Crime Scene	NEUTRAL	h	Room	INDOOR	124.75
25	SURVEY	CS10	Crime Scene	NEUTRAL	h	Room	INDOOR	93.42
26	SURVEY	CS11	Crime Scene	NEUTRAL	h	Room	INDOOR	91.86
27	SURVEY	CS12	Crime Scene	NEUTRAL	h	Kitchen	INDOOR	105.95
28	SURVEY	CS13	Crime Scene	NEUTRAL	h	Room	INDOOR	105.43

29	SURVEY	CS14	Crime Scene	NEUTRAL	h	Bedroom	INDOOR	105.14
30	SURVEY	CS15	Crime Scene	NEUTRAL	h	Bedroom	INDOOR	106.55
31	NAPS	NN1	Natural	NEGATIVE	h	Glass	INDOOR	46.31
32	NAPS	NN2	Natural Scene	NEGATIVE	v	Flooded Building	OUTDOOR	97.23
33	NAPS	NN3	Natural Scene	NEGATIVE	h	Dustbin	INDOOR	95.65
34	NAPS	NN4	Natural Scene	NEGATIVE	h	Broken Bottle	OUTDOOR	137.27
35	NAPS	NN5	Natural Scene	NEGATIVE	h	Toy	INDOOR	108.73
36	NAPS	NN6	Natural Scene	NEGATIVE	h	Skeletons	OUTDOOR	54.45
37	IAPS	NN8	Natural Scene	NEGATIVE	h	Garbage	OUTDOOR	123.39
38	IAPS	NN10	Natural Scene	NEGATIVE	h	CarAccident	OUTDOOR	86.83
39	IAPS	NN11	Natural Scene	NEGATIVE	h	CarAccident	OUTDOOR	69.57
40	NAPS	NN12	Natural Scene	NEGATIVE	h	Crashed Truck	OUTDOOR	124.77
41	NAPS	NN13	Natural Scene	NEGATIVE	h	Bottle	OUTDOOR	86.69
42	NAPS	NN14	Natural Scene	NEGATIVE	h	Skeletons	OUTDOOR	91.29
43	NAPS	NN15	Natural Scene	NEGATIVE	h	Devastated House	OUTDOOR	76.31
44	IAPS	NS2	Natural Scene	NEGATIVE	h	Cemetery	OUTDOOR	90.81
45	IAPS	NS3	Natural Scene	NEGATIVE	h	Dishes	INDOOR	124.33
46	IAPS	NS1	Natural Scene	NEUTRAL	h	Bed	INDOOR	77.94
47	NAPS	NS4	Natural Scene	NEUTRAL	v	Devastated House	OUTDOOR	66.73
48	NAPS	NS5	Natural Scene	NEUTRAL	h	Devastated House	INDOOR	112.43
49	NAPS	NS6	Natural Scene	NEUTRAL	h	Balcony	OUTDOOR	79.24
50	NAPS	NS7	Natural Scene	NEUTRAL	h	Plants	OUTDOOR	101.20
51	NAPS	NS8	Natural Scene	NEUTRAL	h	Water	OUTDOOR	90.46
52	NAPS	NS9	Natural Scene	NEUTRAL	h	Toy	OUTDOOR	103.56
53	NAPS	NS10	Natural Scene	NEUTRAL	h	Objects	INDOOR	151.50
54	NAPS	NS11	Natural Scene	NEUTRAL	h	Bottles	INDOOR	89.91
55	NAPS	NS12	Natural Scene	NEUTRAL	h	Bottle	INDOOR	113.72
56	NAPS	NS13	Natural Scene	NEUTRAL	h	Dishes	INDOOR	123.13
57	NAPS	NS14	Natural Scene	NEUTRAL	h	Bicycles	OUTDOOR	117.08
58	NAPS	NS15	Natural Scene	NEUTRAL	h	Highway	OUTDOOR	76.46
59	NAPS	NN7	Natural Scene	NEUTRAL	h	Crashed Trains	OUTDOOR	96.43

60	NAPS	NN9	Natural Scene	NEUTRAL	h	Boots	OUTDOOR	81.62
----	------	-----	---------------	---------	---	-------	---------	-------