The detection of concealed firearm carrying through CCTV: the role of affect recognition

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfillment of the requirements for the award of Doctor of Philosophy of Loughborough University.

Metadata Record: https://dspace.lboro.ac.uk/2134/9102

Publisher: © Anastassia Blechko

Please cite the published version.
This item was submitted to Loughborough's Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
To my parents
This research aimed to explore whether the recognition of offenders with a concealed firearm by a human operator might be based on the recognition of affective (negative) state derived from non-verbal behaviour that is accessible from CCTV images. Since a firearm is concealed, it has been assumed that human observers would respond to subtle cues which individuals inherently produce whilst carrying a hidden firearm. These cues are believed to be reflected in the body language of those carrying firearms and might be apprehended by observers at a conscious or subconscious level. Another hypothesis is that the ability to recognize the carrier of concealed firearm in the CCTV footage might be affected by other factors, such as the skills in decoding an affective state of others and the viewpoint of observation of the surveillance targets. In order to give a theoretical and experimental basis for these hypotheses the first objective was to examine the extant literature to determine what is known about recognition of affect from non-verbal cues (e.g. facial expressions and body movement), and how it can be applied to the detection of human mal-intent. A second objective was to explore this subject in relation to the detection of concealed firearm carrying through performing a number of experimental studies. The studies employed experts, i.e. CCTV–operators and mainly the lay people as participants. Also, various experimental techniques such as questionnaires and eye-tracking registration were used to investigate the topic.

The results show that human observers seem to use visual indicators of affective state of surveillance targets to make a decision whether or not the individuals are carrying a concealed firearm. The most prominent cues were face, and upper body of surveillance targets, gait, posture and arm movements. The test of decoding ability did not show sufficient relationship with the ability to detect a concealed firearm bearer. The performance on the task might be view dependent. Further research into this topic will be needed to generate strategies that would support reliable detection of concealed firearm carrying through employing of related affective behavioural cues.

**Keywords:** CCTV, gun crime, human affect recognition, non-verbal behaviour
ACKNOWLEDGMENTS

I would like to extend my gratitude to Professor Alastair Gale for giving me a chance and for creating the opportunity to join his group. His expert skills, human insight and humour were invaluable. Special thanks go to Dr Iain Darker who was involved with all aspects of supervision. The progress of the thesis owes a great deal to his valuable comments, criticism and encouragements. He taught me not only how to write academic papers, but also how to program, and how to think through my problems. I am very grateful to both supervisors for bringing out the good ideas in me and for being there to meet and discuss those.

Kevin Purdy, Sarah Love and Ruth Spence provided unobtrusive but effective support, suggestions and encouragement throughout the working on the thesis. I also would like to thank all people of Applied Vision Research Centre, Ergonomics & Safety Research Institute and the Department of Computer Science; all of them were very friendly and helpful.

I would like to acknowledge and thank the CCTV operators and their managers from Charnwood Borough Council and CCTV control rooms in Nottingham and Liverpool for taking the time to explain and to demonstrate their work. Richard Buxton from Mathematical Sciences provided me with a great help on statistical analyses. I am very grateful to Rolf Zon, for his competent support, professional advice and for explaining me the ‘What, Why and How’ of the research world’s casualties and opportunities. Thanks to Dr Moi Jaap for her friendly, joyful and strong personality and for sharing her thoughts and ideas with me. I appreciated her freely and willingly advices. I am grateful to Elena Jakimova and Peter Melis, Katia Kondratieva and her husband Vladimir Kosenko for their pleasant and warm support and in particular for their listening and thinking together about the subject of my work. They also provided me with much needed hospitality and respite for which I am grateful. I would also like to thank Mrs Klaasje Moleman-Baardse for offering her valuable support.

I am particularly indebted to Dirk-Jan Moleman for always being there for me to share the good and bad moments, and for creating the wonderful opportunity to spend three years living on the top of the hill in the most beautiful place on earth, unforgettable Flower Lillies House. It has always been my dream.

It is true to say that this work would never have been done and written without the practical and moral support and encouragements of my dearest parents, Nadia and Victor Blechko, very special people.

Мои дорогие папа и мама, спасибо вам за радость быть вашей дочерью, спасибо за мысли и тревоги обо мне, спасибо за ваше тепло и заботу, без которых я не смогла бы осуществить того, что осуществила.
# CONTENTS

CHAPTER 1: INTRODUCTION .................................................................................................. 1

1.1. Background context ........................................................................................................ 1
   1.1.1. MEDUSA project ................................................................................................... 2
   1.1.2. CCTV ..................................................................................................................... 3
   1.1.3. CCTV operator tasks .............................................................................................. 6

1.2. Exposure of criminal’s emotional state during or before committing a crime ............... 11

1.3. Recognition of affective state from human appearance .................................................. 14
   1.3.1. Recognition of affect from facial expressions and human motion ....................... 15
   1.3.2. Viewpoint dependence in perception of human gender, actions and affect .......... 22

1.4. The present research ..................................................................................................... 24

1.5. Thesis structure ............................................................................................................. 27

CHAPTER 2: METHODOLOGY. ............................................................................................. 30

2.1. Introduction .................................................................................................................. 30

2.2. Survey of CCTV control rooms .................................................................................... 32

2.3. Creating the mock CCTV footage of people carrying firearms .................................... 34

2.4. Measure the performance on the task of recognition a concealed firearm carrier by observers: Signal Detection Theory .................................................. 41

2.5. Measurement of affective state ..................................................................................... 43
   2.5.1. Aggression Questionnaire ..................................................................................... 43
   2.5.2. Multiple Affect Adjective Check List- Revised (MAACL-R; state and trait form) ............................................................................................................................... 44

2.6. Measurement of sensitivity to non-verbal cues of human observers: Profile of Nonverbal Sensitivity (PONS) test ................................................................. 45

2.7. Detection of the visual cues related to concealed firearm carrying ................................ 48
   2.7.1. Cue-detection questionnaires. Questionnaire design ............................................ 48
   2.7.2. Eye tracking method ............................................................................................... 50
   2.7.3. Mock CCTV footage used in the experiment with eye movements registration ................................................................................................................. 52

2.8. Conclusion .................................................................................................................... 53
5.3. Discussion ......................................................................................................................... 95
5.4. Conclusion ........................................................................................................................ 98

CHAPTER SIX: EXPERIMENT FOUR ................................................................................... 100
6.1. Introduction ..................................................................................................................... 100
6.2. Method ............................................................................................................................ 103
    6.2.1. Participants ............................................................................................................ 103
    6.2.2. Materials ............................................................................................................. 103
    6.2.3. Procedure ............................................................................................................. 106
6.3. Results ............................................................................................................................ 106
6.4. Discussion ....................................................................................................................... 108
6.5. Conclusion ...................................................................................................................... 110

CHAPTER 7: EXPERIMENT FIVE AND SIX .......................................................................... 112
7.1. Introduction ..................................................................................................................... 112
7.2. Experiment Five .............................................................................................................. 115
    7.2.1. Method .................................................................................................................. 115
        7.2.1.1. Participants ..................................................................................................... 115
        7.2.1.2. Materials .......................................................................................................... 116
        7.2.1.3. Procedure ........................................................................................................ 119
    7.2.2. Results .................................................................................................................. 119
    7.2.3. Discussion ............................................................................................................. 123
7.3. Experiment Six ................................................................................................................ 124
    7.3.1. Method .................................................................................................................. 124
        7.3.1.1. Participants ..................................................................................................... 124
        7.3.1.2. Materials .......................................................................................................... 124
        7.3.1.3. Procedure ........................................................................................................ 125
    7.3.2. Results .................................................................................................................. 126
    7.3.3. Discussion ............................................................................................................. 131
7.4. Conclusion ...................................................................................................................... 132
LIST OF FIGURES

FIGURE 1.1: CCTV OPERATORS MONITOR CCTV SCREENS IN THE CONTROL ROOM IN NOTTINGHAM, THE UK (PHOTOGRAPH TAKEN BY THE AUTHOR) .................................. 9

FIGURE 1.2: COGNITIVE REPRESENTATION OF AFFECT ACCORDING TO RUSSEL (1980) WHICH CONSISTS OF TWO MAIN DIMENSIONS (MISERY VERSUS PLEASURE AND AROUSAL VERSUS SLEEPINESS) AND FOUR REMAINING VARIABLES WHICH DO NOT FORM INDEPENDENT DIMENSIONS. ........................................................................................................................................ 17

FIGURE 1.3: EXAMPLE OF VISUAL PRESENTATION OF POINT-LIGHT DISPLAYS IN PROFILE VIEW (ADAPTED FROM JOHANSSON (1973)) ........................................ 18

FIGURE 2.1: SCHEMATIC REPRESENTATION OF THE FILMING SET-UP .................. 37

FIGURE 2.2: EXAMPLES OF CLOTHING OUTFIT USED IN THE EXPERIMENT ............... 37

FIGURE 2.3: PRESENTED FIREARMS: A - .32 CALIBER REVOLVER; B – GLOCK; C- SAWN-OFF SHOTGUN ........................................................................................................... 40

FIGURE 2.4: PRESENTED MATCHED OBJECTS – BOTTLES OF DIFFERENT SIZES AND WEIGHTS .............................................................................................................. 41

FIGURE 2.5: SET-UP OF THE EYE-TRACKING EXPERIMENT. .................................. 51

FIGURE 4.1: EXAMPLES OF STILL IMAGES OF THE IDEALISED FOOTAGE (A- WALKING WITH A CONCEALED FIREARM; B- WALKING WITH A CONCEALED TWO-LITRE BOTTLE); (ADAPTED FROM BLECHKO ET AL. (2008)) .................... 69

FIGURE 4.2: SENSITIVITY ($da$) TO A CONCEALED FIREARM AMONGST CCTV OPERATORS AND LAY PEOPLE (ADAPTED FROM BLECHKO ET AL. (2008)) .......... 74

FIGURE 4.3: CORRELATION BETWEEN THE SIZE OF THE INFLUENCE OF CARRYING A FIREARM ON THE SURVEILLANCE TARGET’S ANXIETY LEVEL AND THE NUMBER OF TIMES THAT THE SURVEILLANCE TARGET WAS DEEMED TO BE CARRYING A CONCEALED FIREARM; (ADAPTED FROM BLECHKO ET AL., 2008)) ........................................................................................................ 75
FIGURE 5.1: MEAN SCORES ON THE DIFFERENT SCALES OF MAACL-R QUESTIONNAIRE (ANXIETY, DEPRESSION, HOSTILITY, DYSPHORIA, POSITIVE AFFECT, SENSATION SEEKING AND PASS) IN INNOCUOUS OBJECT (I.E., SURVEILLANCE TARGET WITH A BOTTLE) AND GUN (I.E., SURVEILLANCE TARGET WITH A FIREARM) CONDITIONS, WHICH WERE GIVEN BY OBSERVERS IN THEIR RATING OF THE SURVEILLANCE TARGETS’ AFFECTIVE STATE. ......................... 92

FIGURE 6.1: THE SURVEILLANCE TARGETS’ AFFECTIVE STATE ESTIMATED BY OBSERVERS. THE SURVEILLANCE TARGETS’ AFFECTIVE STATE IS REPRESENTED BY THE MEAN SCORES ON THE SCALES OF MAACL-R IN INNOCUOUS OBJECT (I.E., CARRYING A BOTTLE) AND GUN (I.E., CARRYING A FIREARM) CONDITIONS; (ADAPTED FROM BLECHKO ET AL., 2009))........................................... 107

FIGURE 7.1: EXAMPLE OF SMALLEST (A) AND LARGEST (B) FRAMES FROM A VIDEO SEQUENCE, DEMONSTRATING EYE FIXATIONS AND THREE AOIs (I.E., FACE, UPPER BODY, LOWER BODY) ANNOTATED WITH DIMENSIONS IN CM AND DEGREES OF VISUAL ANGLE FOR HEIGHT (H) AND WIDTH (W) OF EACH AOI ....................................................................................................................................... 117

FIGURE 7.2: THE SURVEILLANCE TARGETS’ AFFECTIVE STATE, ESTIMATED BY OBSERVERS. THE SURVEILLANCE TARGETS’ AFFECTIVE STATE IS REPRESENTED BY THE MEAN SCORES ON THE SCALES OF MAACL-R IN INNOCUOUS OBJECT (I.E., CARRYING A BOTTLE) AND GUN (I.E., CARRYING A FIREARM) CONDITIONS...................................................................................................... 120


FIGURE 7.5: SENSITIVITY (da) TO A CONCEALED FIREARM AMONGST TWO GROUPS OF PARTICIPANTS (I.E., WHO RECEIVED INSTRUCTIONS WITH DESCRIBING OF AFFECTIVE CUES AND WHO RECEIVED INSTRUCTIONS WITHOUT THIS DESCRIBING. ................................................................. 127

FIGURE 7.6: COMPARISON OF THE EFFECT OF TYPE OF INSTRUCTION (WITH AFFECTIVE CUES VERSUS WITHOUT AFFECTIVE CUES) ON MEAN DWELL TIME OF OBSERVERS WITHIN THREE AOIs (I.E., FACE, BODY, AND LEGS) ......................... 129
FIGURE 7.7: COMPARISON OF THE EFFECT OF TYPE OF OBJECT (A - GUN VERSUS B – INNOCUOUS OBJECT) ON MEAN DWELL TIME OF OBSERVERS WITHIN THREE AOIs (I.E., FACE, BODY, AND LEGS). ........................................................ 130
LIST OF TABLES

TABLE 3.1: OVERALL NUMBER OF THE CLIPS THAT WERE FILMED ACROSS ALL PARTICIPANTS ....................................................................................................................... 57

TABLE 3.2: OVERALL NUMBER OF CLIPS MADE WITH EVERY KIND OF CARRYING OBJECT (I.E., THREE TYPES OF BOTTLES AND THREE TYPES OF FIREARMS) IN EACH SESSION, FOR ONE PARTICIPANT. ................................................................. 58

TABLE 3.3: MAACL-R SCORES ON DYSPHORIA, ANXIETY, DEPRESSION AND HOSTILITY SCALES FOR TWO CONDITIONS: GUN AND INNOCUOUS OBJECT ............... 62

TABLE 4.1: SUMMARY SCORES FOR CCTV OPERATORS AND LAY PEOPLE ON PONS TEST (I.E., PROPORTION CORRECT ANSWERS); (ADAPTED FROM BLECHKO ET AL.(2008)) ........................................................................................................ 76

TABLE 5.1: OVERALL NUMBER OF CLIPS USED IN EVERY TRIAL (INNOCUOUS OBJECT AND GUN) WITH DIFFERENT KIND OF CARRYING OBJECT (I.E., THREE TYPES OF BOTTLES AND THREE TYPES OF FIREARMS), FOR EACH SURVEILLANCE TARGET ...................................................................................................... 89

TABLE 5.2: COMPARISON OF MEAN SCORES OF OBSERVERS ON ALL SCALES AND SUBSCALES OF MAACL-R QUESTIONNAIRE IN THE INNOCUOUS OBJECT AND GUN CONDITIONS (REPEATED MEASURES ANOVA WITH df = 1/30). *p<.05 .......... 91

TABLE 5.3: COMPARISON OF MEAN DIFFERENCES IN SCORES (%) OF OBSERVERS AND SURVEILLANCE TARGETS ON ALL SCALES AND SUBSCALES OF MAACL-R QUESTIONNAIRE IN THE INNOCUOUS OBJECT AND GUN CONDITIONS (REPEATED MEASURES ANOVA WITH df = 1/30). *p<.05 ......................... 93

TABLE 6.1: OVERALL NUMBER OF CLIPS USED IN THE STUDY IN EACH KIND OF TRIAL (INNOCUOUS OBJECT AND GUN), FOR EACH SURVEILLANCE TARGET .......... 105

TABLE 7.1: OVERALL NUMBER OF CLIPS USED IN THE STUDY IN EACH KIND OF TRIAL (INNOCUOUS OBJECT AND GUN), WITH DIFFERENT KIND OF CARRYING OBJECT (I.E., BOTTLE AND FIREARM), FOR EACH SURVEILLANCE TARGET .......... 118
LIST OF APPENDICES

AGGERESSION QUESTIONNAIRE (BUSS & PERRY, 1992). ........................................APPENDIX 1
MAACL-R STATE QUESTIONNAIRE (LUBIN & ZUCKERMAN, 1999). ......................APPENDIX 2
MAACL-R TRAIT QUESTIONNAIRE (LUBIN & ZUCKERMAN, 1999). ......................APPENDIX 3
CUE-DETECTION QUESTIONNAIRE ..........................................................APPENDIX 4
PARTICIPANT INFORMATION SHEET ........................................................APPENDIX 5
INFORMED CONSENT FORM .................................................................APPENDIX 6
PARTICIPANTS’ INSTRUCTIONS FOR EXPERIMENT ‘MEDUSA: Eye-movements associated with recognition of mood in people observed via CCTV’. APPENDIX 7
DEBRIEFING SHEET ...........................................................................APPENDIX 8
CHAPTER 1: INTRODUCTION

1.1. Background context

Illegal gun possession and the use of illicit firearms is an example of serious crime which is of major importance these days. The late 1990s and early 2000s saw a major increase in the levels of gun crime within the UK (Povey, Coleman, Kaiza, Hoare, and Jansson, 2008). This surge in gun crime was accompanied by a change in the types of people carrying and using illicit firearms: “Illegal weapons were thought to be less and less the exclusive preserve of the ‘professional’ career criminal and, instead, were falling into the hands of a young, amoral and dangerously unstable class of offenders for whom firearms meant power and respect” (Squires, 2000, p. 98). More recently the number of inexperienced, ‘amateur’ criminals who possess, carry and use firearms is growing, as suggested by an increase in non-fatal gun crime in the UK and the outcomes of interviews with senior police officers (Lockhart, McClory, and Qvortrup, 2007). The available crime statistical data show that firearm offences occur in a number of major cities in the UK (Booth, Black, Dear, John, Johnson, Levy, Selvan, and Williamson, 2008; Squires, Silvestri, Solomon, and Grimshaw, 2008). It has been noted that the murder became increasingly ‘weaponised’ (knives and firearms), especially amongst younger people (Booth et al., 2008). In London, for example, between 2005 and 2007 weapons, in general, were used in 91% of life threatening incidents, and in 12 cases (i.e. 38% of all of these incidents) the murder weapon was a firearm. Some 12% of all gun crimes committed in England and Wales took place in Manchester, where 1,240 firearm offences were recorded in the year 2002-2003. Recorded violent crime (including firearm offences) in the West Midlands has risen continuously from 2003 to 2007. Moreover, in Liverpool there was an overall increase in the number of total offences (including firearm offences) between 2002-2003 and 2006-2007. For the period of one year (2006-2007) in 42% of
the total recorded firearm offences (i.e. 172 from 410 incidents) there was firearm-involved violence against the person, including murders, attempted murders and other acts (Squires et al, 2008). Firearms have been linked to the activities of drug gangs and some reports indicate that firearm offences have dropped slightly in very recent years possibly as a result of successful police action against these gangs. However, gun crime remains a major ongoing concern.

The growth of gun crime and other potential gun related incidents led to an EPSRC Ideas Factory Sandpit in 2005 on the prevention of gun crime which was hosted by the EPSRC, the Home Office and the Metropolitan Police. This was attended by a group of senior experienced research scientists, many of whom had no experience of gun crime, and examined a wide range of gun related research ideas seeking to develop new research methods and techniques in the domain. Proposed research was informed by those research colleagues attending who had research experience related both to gun crime and criminology. The event was carried out under the steer and supervision of a panel of experts from the police and defence establishments. The outcome from the Sandpit was a small number of EPSRC funded research projects which sought to examine gun crime from a wide variety of different perspectives. One of these projects was MEDUSA (http://www.lboro.ac.uk/research/applied-vision/projects/medusa/index.htm).

1.1.1. MEDUSA project

MEDUSA (Multi Environment Deployable Universal Software Application) was a project proposed by Professor Alastair Gale and involved five University research departments (Loughborough’s Applied Vision Research Centre, Kingston’s Digital Imaging Research Centre, Central Lancashire’s Applied Digital Signal and Image Processing Research Centre, Liverpool’s Veterinary Clinical Science Department and Brighton’s School of Applied Social Sciences). The project was overseen by a steering group comprising: ACPO (Association of Chief Police Officers) CCTV/ Video Working Group; ACPO Criminal Use of
Firearms; CCTV (Closed Circuit Television) User Group; Crime and Intelligence Technical Support, Metropolitan Police; LGC Forensics, and the National Firearms Centre, Royal Armouries. MEDUSA was concerned with the identification of those situations associated with gun related threats, based on the behavioural interpretation of CCTV data. The project aimed to develop a new machine learning system for the detection of individuals carrying guns through the use of CCTV surveillance networks. In doing this the system combined both psychological and machine vision approaches. The basis of the project was that currently extensive CCTV networks exist in most cities where many cameras feed into control rooms where the operators can only monitor a small selection of camera inputs at any one time. The concept was to develop software which would automatically detect potential gun related incidents on any of the camera inputs and then prioritise the CCTV feed from that camera to the human operator in the control room along with an alert signal. CCTV operators are tasked with monitoring and trying to detect a wide range of incidents. Potential gun crime, despite its reported growth in incidence, still represents a very rare possible daily event which in reality would most probably not be detected by a human operator. A key approach in the project was to examine what factors underpinned how humans identified guns in CCTV images as a basis of both informing and benchmarking the development of suitable machine vision approaches. The research examined many aspects of CCTV and gun crime activity. Real CCTV footage of gun related incidents were acquired nationally and studied. This led to the establishment of a set of carefully set up and staged CCTV footage of individuals walking whilst they sometimes carried guns or gun sized objects either overtly or covertly. This footage was used extensively in the MEDUSA project. The research presented here was carried out as part of the MEDUSA project.

1.1.2. CCTV

In the UK, the use of CCTV for public space surveillance burgeoned in the 1990s, largely under the auspices of crime reduction (Norris and Armstrong,
CCTV is a system in which a number of video cameras are connected in a closed circuit or loop, producing images which are sent to a central television monitor or recorded (Goold, 2004), one of the primary utilities of which is to prevent and reduce crime as the cameras can be used in order to assist police in the timely detection and arrest of offenders identified by CCTV operators. Visual surveillance via CCTV systems involves the use of cameras and recording systems in order to monitor the cameras for certain events and incidents which may include “the presence or behaviour of people, changes in process, anomalies in expected conditions, verification of standard operating conditions or protocols, or the detection of specific threats or circumstances (Donald, 2005, p.3). Contemporary CCTV networks provide extended surveillance of public spaces in town and city centres, outside or inside of the buildings (e.g., shops), and on housing estates (Gill, Spriggs, Allen, Hemming, Jessiman, Kara, Kilworth, Little, Swain, 2005) that can possibly aid crime reduction. Thus, potentially, CCTV networks could allow the targeting of different aspects of crime. Indeed, some studies which compared a number of large towns without CCTV centre with towns with CCTV surveillance showed that the surveillance system did increase detection rates for human violence (Sivarajasingam, Shepherd, Matthews, 2003). Another study demonstrated that the installation of city centre CCTV was followed by increased detection of violence and disorder by police, most likely in collaboration with CCTV operators (Sivarajasingam, Shepherd, Walker, Walters, and Morgan, 2005).

However, various other studies, reviewed and performed by Gill and Spriggs (2005) (e.g., Brown, 1995; Deismann, 2003; Welsh and Farrington, 2002) suggest that CCTV generally fails to reduce overall crime rates significantly. Welsh and Farrington (2002) found dissimilar effects of CCTV on different kind of crime. Their review showed that there was no impact of CCTV on crimes of violence, but they found a significant effect on vehicle crime. Furthermore, personal crimes (e.g. assault) were less likely to be influenced by CCTV (Deismann, 2003), but property crimes were found to be more susceptible to the impact of CCTV (Brown, 1995). The main purpose of the data collection and analysis of the study of Gill and Spriggs (2005) was to
measure an impact of the CCTV projects on a number of outcomes such as for example the change in crime level and fear of crime as results of installation of CCTV. The results showed that the CCTV had little overall effect on crime levels. Only a very few changes in crime levels were larger than could be due to chance alone.

With regard to gun crime, there is too little evidence relating to the impact of CCTV on such crime (Darker, Gale, Ward, and Blechko, 2007). The analysis of all CCTV incidents in two years (i.e., between March 2006 and February 2008) performed at a CCTV control room of a large UK city (mid-year population estimate, 2006: 300,000 to one significant figure) showed that only 10 of the 2253 incidents spotted by operators through proactive monitoring involved firearms (Darker, Gale, Blechko, 2008). Nowadays, CCTV is used mostly in order to co-ordinate a response to an incident, along with gathering evidence that can be used to direct investigations (Brown, 1995). Although training guidelines for CCTV operators stress the proactive use of CCTV (Diffley & Wallace, 1998), it appears that it is being used mostly as a reactive or retrospective tool to tackle crime.

In the context of gun crime, the detection of people carrying firearms and removing those weapons before they could be used would be the ideal proactive solution. Current CCTV surveillance practices are known to be able to predict and, in the course of proactive monitoring, to spot mal-intent by detecting the precursors of overtly violent behaviour, as shown by a number of researchers (e.g., Norris and Armstrong, 1999; Troscianko, Holmes, and Stillman, 2004). Norris and Armstrong (1999) noted that CCTV operators proactively searched local hot spots for criminals based on their instinct and experience. They described a real-life example of proactive surveillance by a CCTV operator that was effective. In addition, they observed situations in which human operators proactively spotted suspicious behavior and alerted the police. Some of those cases involved the detection of illegal firearms. Furthermore, the study of Troscianko et al. (2004) showed that the prediction of lawless behavior via CCTV can be performed accurately by human observers. The examples of situations in which CCTV operators have been able to detect
firearms in the course of the proactive monitoring of CCTV networks have also been documented in the study of Darker et al. (2007).

A major cause of a possible failure of a CCTV system to detect mal-intent lies in the fact that there are many factors which are known to have an effect on the effectiveness of the surveillance activities of CCTV operators. Such technology-related failures as the poor configuration and positioning of CCTV cameras and insufficient quality of video recording have been demonstrated and discussed in several studies (e.g., Gill and Spriggs, 2005; Gill et al., 2005; Luff, Heath, and Jirotka, 2000; Keval and Sasse, 2010). Besides the failures of technology, there are also a number of other, human-related factors that restrict the utility of CCTV in the detection of mal-intent in general and mal-intent involving firearms in particular. It will be discussed in the next section in the context of the tasks of surveillance operators.

1.1.3. CCTV Operator Tasks

From the previous published research (e.g., Wallace and Diffley, 1998) more understanding has been gained of the way that CCTV operators work and are being trained, and particularly of their methods of spotting suspicious human behaviour. The work of CCTV operators consists of miscellaneous activities that may cause interference to each other. Although the work of CCTV operators mainly includes monitoring incidents, detection, recognition and identification (Wallace and Diffley, 1998; Gill et al., 2005), various activities other than monitoring are part of the job. While monitoring images from CCTV cameras, the operators are also in constant communication with guards, police and supervisors, which is one of the essential aspects of their job. This communication implies contact with the guards or police who should be informed by operators and, if necessary, directed toward the suspicious individuals in those situations when suspicious individuals are identified from the CCTV images. Sometimes the operators would give an exact description of that person or contact the police in order to know the identity of a car driver. Regularly a situation can occur when operators need to respond to the police
request to focus a camera on a certain place, so instead of being pro-active the operators often act reactively. Furthermore, in case if a problem occurs the operators need to communicate about it with their supervisors in order to provide back-up. Besides that, a lot of communication takes place between operators and their management due to various situations.

Another human-related factor which may affect crime detection performance is overload of visual information that should be processed by operators during their monitoring activities. A single operator is usually required to monitor the images derived from a large number of cameras (Gill et al., 2005). Consequently, the capacity of CCTV operators to perform continuous live observation of the CCTV images is limited due to the fact that their vigilance might be impaired. The vigilance of human operators is known to be maintained effectively during only a limited amount of time (about an hour, according to Tickner and Poulton, 1973), where after a ‘visual neglect’ or a misinterpretation of visual information can occur. In relation to the statistics of gun crime detection, it can be assumed that since the incidents involving firearms are spotted through CCTV with relative rarity in comparison to other events, it restricts the possibility for CCTV operators to maintain their vigilance to this event and to learn the indicators of potential gun crime (Darker et al., 2007). Along with other human-related and technology-related factors, mentioned above, decreased human vigilance can prevent the operators from reliable detection and identification of criminals’ activities in general, and the gun crime in particular, and can make this task extremely difficult.

Moreover, there is another factor that should not be forgotten in relation to the effectiveness of operators’ performance - work experience and personal qualities of workers. In this regard, Donald (1999) emphasized that the identification and the selection of the right people for the job of a CCTV operator, along with the right training, are essential for effective CCTV system functioning. Various published studies provide guidance on the selection and recruitment of CCTV operators. A CCTV operator should know when and where to look through constant (and random) monitoring of images along with using the alarms from Police or local authorities, whereby operators’ past
experience, training and culture are very important (Aldridge, 1994). It is known that the abilities, knowledge, skills and personal qualities of candidates can be described by means of different competencies (Diffley and Wallace, 1998; Wallace and Diffley, 1998). According to Wallace and Diffley (1998) “a competency is something a person can demonstrate, i.e. a behaviour that proves they are able to perform a task or activity” (p. 4). Having knowledge of the competencies of a candidate is necessary to establish clear training needs, although there seems to be a mutual dependence between competence and training. Diffley & Wallace (1998) point out that training is required to add necessary knowledge and skills to existing competence, creating a new, more developed competence, even though training cannot provide an operator with full competence.

Through dialogue with operators and their managers (it will be discussed in more details in Chapter 2) it has been found that they think that good performance amongst CCTV operators depends not only on proper training but also on the ‘innate abilities’ of candidates, although there is still little research to support this assertion (e.g., Troscianko, Holmes, and Stillman, 2004; Darker, Gale, Blechko, and Whittle, 2009). It has also been stated that people may have certain qualities that make them able to show good performance in the monitoring task irrespective of any training. Thus, while a comprehension of how to teach people to spot suspicious behaviour from CCTV imagery adds a significant contribution to knowledge about the best way to enhance operators’ performance, it does not give a complete picture. To make the selection and training of CCTV operators successful it is necessary to identify individual differences in specific skills, and to decide to what extent they need to be trained.

As mentioned earlier, the examination of the tasks of surveillance operators has shown that the most important purpose of their job is to detect and respond to incidents or signals that threaten the safety of people by means of visual observation of monitors (Gill et al., 2005; see Figure 1.1). The operational requirements check list provided by the CCTV Operational Requirements Manual (Aldridge, 1994) describes four task categories:
monitoring and control (*i.e.* determine the number, direction and speed of movement of people in the images); detection (*i.e.* ascertain whether or not a person is visible in the images); recognition (*i.e.* to be able to say whether or not the individual shown is the same as someone who had been seen before); identification (*i.e.* establishing the identity of a subject). Gill *et al.* (2005) presented data collected from thirteen control rooms in England which demonstrated that operators can spend up to 58 per cent of their working time monitoring incidents. Because the monitoring incidents employs more than half of the working time of CCTV operators, the visual information analysis is one of the core competencies required by CCTV operators (Wallace and Diffley, 1998, Donald, 2001).

![CCTV operators monitor CCTV screens in the control room in Nottingham, the UK](image)

*Figure 1.1. CCTV operators monitor CCTV screens in the control room in Nottingham, the UK (photograph taken by the author)*

This implies that operators need to have the skills to make sense of visual information and to recognize when something is wrong or is likely to go wrong from information viewed on the monitors. The use of visual information (*i.e.* visual cues) is very important in potential proactive gun crime detection. The study of Darker *et al.* (2007) showed that the gun crime had been detected proactively, although the CCTV operator was monitoring the event for a reason other than the suspicion of gun crime. In each observation a gun was detected only when the gun was in view or when the shape of the gun was readily detectable. Thus, in the situation when the gun is being carried in an
unconcealed manner, the shape of the gun itself is the first visual cue that gives an indication of a potential future gun crime. Although this study revealed only the cues that CCTV operators were aware of using and which they could verbalise (i.e. overt cues), it was also posited that the perceptual elements in the task of recognition of human intentions to commit gun crime may tend to operate at a subconscious level as well. That means that also covert visual cues could be used by human operators in this kind of task. Since carrying a firearm is not necessarily accompanied by overtly violent behaviour and the firearm may be concealed to some degree or completely (Hales, Lewis, and Silverstone, 2006), the perception of covert cues (e.g., specific, subtle behavioural patterns related to concealed firearm carrying) might be even more important.

As the hidden firearm is considerably less visible to observers than the firearm carried in plain view, the detection of firearms concealed underneath a person's clothing is an important obstacle to the improvement of the security of the general public as well as the safety of public places. Accordingly, it gives rise to a question about how well and accurately human operators would be able to detect a carrier of a concealed firearm through the observation of CCTV imagery. The fact that firearms may be hidden makes the recognition of the firearm bearer much more difficult, and perhaps it is highly dependent on such factors as individual abilities of human surveillance operators, appropriate training and experience. One of the important individual characteristics which underlies the skill to recognize crime involving firearms through CCTV, and which has not yet been investigated in the context of CCTV surveillance, is the ability of a CCTV operator to identify the behaviours that indicate that someone is carrying a concealed firearm through observation of non-verbal cues associated with firearm carriers’ affective state and made from different viewpoints. According to operators, their managers and researchers (Donald, 2006) the understanding of body language by a CCTV operator is crucial, and behaviour based techniques focusing on body language are required in the CCTV monitoring industry. In order to provide more extensive support for the assumption that the interpretation of the intentions of others through non-verbal
behaviour and the ability to decode emotional state through observation of non-verbal cues might be essential for the performance of CCTV operators on the task of recognising a concealed firearm carrier, the next sections will discuss the evidence from applied research suggesting that the act of committing a crime is indeed associated with a certain affective response. In addition, the issues surrounding the ability to decode affective state through observation of non-verbal cues will be discussed, along with another influential factor in perception of human motion, postures and affect such as viewpoint dependence.

1.2. Exposure of criminal's emotional state during or before committing a crime

Several studies found evidence that during a crime, or by preparing oneself to commit a crime, some affective processes are more likely to be present and to be reported by offenders than others, for example feelings of excitement, anxiety or anger (Cusson, 1993; Canter and Ioannou, 2004; Katz, 1988). Canter and Ioannou (2004) investigated the emotions that are experienced during a variety of offences. The hypothesis that during various types of criminal act the offenders experience different emotions was tested. On the basis of responses to interviews with convicted criminals it was found that both positively and negatively loaded emotional responses are linked to different kinds of crimes. The most distinct positive emotion was exhilaration, which was primarily associated with such crimes as robbery and property crimes. Anger was the most distinctive negative emotion which had its highest means for violence and murder.

Fear, another very influential affective response of offenders when committing a crime (Cusson, 1993), has been used to explain the connection between deterrence theory and situational crime prevention. Cusson (1993) emphasizes the influence of the emotional component on criminal behaviours,
namely on their decisions before or during the criminal event. In respect of the offender’s behaviour, deterrence is “the inhibiting influence that fear exercises over the potential offender” (p. 56). Fear can play an important role in steering offenders’ behaviour. The effects of fear are often beyond their control, and can even be sufficient to prevent the commission of a crime. According to Walsh (1986) 46% of all robbers interviewed in his study felt fear just before the robbery and 52% during the execution of the crime. Fear was also the dominant emotional state experienced by muggers during their first mugging (Lejeune, 1977). The affective response of an offender to an offence has been found to be vulnerable to different factors as for example the presence of surveillance, alarms, possible following sanctions (e.g., arrest) etc.. The cognitive appraisal of these factors by an offender cannot only induce a feeling of fear, but can also modify the physical behaviour of a person.

Another influential factor which can cause a change in the criminal’s affective state is the experience of carrying a gun. With respect to firearms offences, the previous research showed that even seeing a firearm or simply identifying weapons was found to be sufficient to increase the accessibility of aggressive thoughts by ordinary people (Anderson, Benjamin, Arlin, and Bartholow, 1998), which is known as the ‘weapons effect’. Klinesmith, Kasser, McAndrew, (2006) investigated whether interacting with a gun (i.e., a toy gun) increased testosterone levels and later aggressive behaviour. According to their results the exposure to guns triggers changes in testosterone levels, which in turn increase interpersonal aggressive behaviour. Berkowitz, LePage (1967) also reported evidence of increased aggressive behaviour of angered participants as reaction on the presence of guns, compared to badminton racquets and shuttlecocks.

Evidence to suggest that offenders tend to undergo different emotional states during unlawful activity that is associated with carrying an illegal firearm has been also found in recent research conducted in England and Wales and commissioned by the Home Office (Hales et al., 2006). This research covered a number of areas where collectively more than half of recorded gun crimes in England and Wales take place. In the report, 12 case studies were presented
which explored four different themes: illegal drug markets, robbery and burglary, gang violence and firearm possession offences. By means of interviews, the authors had collected offenders’ accounts of their own emotional responses to the carrying of firearms. The majority of respondents reported that carrying a gun was associated with feelings of safety and/or empowerment. The minority of interviewed offenders indicated that carrying a firearm made them feel nervous of being arrested or to have to use it. According to one interviewee from London, possessing a gun activates a combination of such emotions as fear and empowerment.

Although, the studies reviewed above provide evidence that carrying an illegal firearm is associated with a range of emotions, which can be consciously appraised and reported by the offender, primarily, the negatively loaded affective states (e.g., aggression and fear) might be seen to be significantly related to committing a gun crime. It can be concluded that although just having a gun concealed does not always mean that an offence will take place immediately, carrying an illegal weapon as a first step towards committing a crime may evoke certain affective states in the would-be offender. This affective state in turn may result in a change of his or her body movement and facial expressions which can potentially be spotted through careful observation of physical signs or emotional cues conveyed by the body and face (Ekman, 2003; Atkinson, Dittrich, Gemmell and Young, 2004). The abovementioned research suggests that the ability to recognize in particular the negative affective state through perception of non-verbal expressions of surveillance targets might enhance the ability of human operators to detect an unlawful act accompanied with using a firearm.

The next section will review a number of psychological and neuropsychological studies which indicate that the intentions of others can be inferred from their emotional state on the basis of the perception of facial expressions and body movement with relatively high levels of accuracy. The previously discussed studies which investigated the emotions that are experienced during offences showed that the recognition of negative affect might be seen to be related to recognition of the firearm carriers. This next
section therefore describes the research on affect recognition in general and negatively loaded affect in particular from facial expressions and the human gait. It will be supported by the evidence from neuropsychological studies on how the human and mammalian brain is involved in the processing of affective information in general and negative affective information in particular.

1.3. Recognition of affective state from human appearance

Theorizing about the recognition of emotions of other people has a history which extends from Charles Darwin’s classic work: *The expression of emotions in man and animals* (1872/1965). In the years following this influential publication it has become apparent that the ability to understand non-verbal cues supports human social interaction (Mehrabian, 1981) and non-verbal cues are an important means in the transmission of feelings, in addition to spoken language (Ekman, 2003). Humans seem to be able to judge others’ emotions accurately at above chance levels based even on small amounts of behavioural information (Hall and Bernieri, 2001). This information about the behaviour of others can span less than 1 second, but more often will span several seconds to several minutes (Ambady and Rosenthal, 1992). The existence of a large body of research on recognition of emotions from facial expressions and body movements shows that it is worthwhile to discuss these studies more extensively. It needs to be mentioned that there is no clear definition of term ‘emotion’ and there can be therefore different interpretations and meanings of this word. It might be understood as a state of conscious feelings characterized by physiological changes such as arousal or as more automatic response which may lack the range and variety of conscious emotion (Baumeister, Vohs, DeWall, and Zhang, 2007). In the present thesis the terms affect and affective state will be used to refer to what is believed to be the automatic processes in response to experience of carrying a firearm.
1.3.1. Recognition of affect from facial expressions and human motion

Examples of classic psychological investigations on the facial expression of emotion lie in the work of Paul Ekman and colleagues, which dates back to the 1960s (Ekman, 2003). By comparing head and body cues, Ekman (1965; Ekman and Friesen, 1967) found that these cues provide different information about apparent emotions to an observer. It was found that the head is more informative about the type of emotion (whether the person feels angry, afraid, sad, etc.), whilst the body is more informative about the intensity of an emotion. The assumption that there are two channels of non-verbal communication has been examined in this early research on the recognition of emotions from facial expressions and bodily movements. On the basis of this research, Meier-Faust (2002) made an assumption that human body language can be divided into two categories: structural information (e.g., facial features, body build, and general appearance) and kinetic information (e.g., facial expressions, gestures, body movements, posture). Structural information can tell us about what kind of emotion a person is experiencing. Conversely, body movements and posture indicate the intensity of emotion and can physically illustrate what someone feels.

Ekman also wanted to understand what information facial expressions give away about a person. Ekman and Friesen (1969) provided evidence that people who suppress emotional states often 'leak' the true emotion in fleeting facial micro-expressions which appear on the face for just 1/25th of a second. It has been theorized that these flashes of emotions are beyond our conscious control and so cannot be masked. Further, they can be spotted by a trained person. They found that human facial expressions are not as accurate and controllable as people may think. Attempts to hide an emotion with a fake expression can therefore be foiled by these micro-expressions which give away the true, underlying emotion.

Although research on facial expression has a long history and a strong foundation, according to some researchers there are shortcomings which call for new research on recognition of facial affect. According to Russell (1994) the
attempt to support the claim that emotions are universally recognized from facial expression by using existent methods is questionable because of its ecological, convergent, and internal validity. Russell claims that emotion in general (and facial expression of emotion in particular) can be best characterized in terms of a multidimensional affect space, rather than discrete emotion categories (such as "fear" or "happiness" as according to Ekman). Further, he assumes that to be able to characterize affect space, just two orthogonal bipolar dimensions of pleasure-misery and arousal-sleepiness (see Figure 1.2.) are sufficient (Russell, 1980). This two-dimensional space can be further elaborated by four other variables: excitement, contentment, depression and distress.

Emotions are not a necessary or sufficient precondition of certain spontaneous expressions according to Fernandez-Dols, Sanchez, Carrera, and Ruiz-Belda (1997), which is not in line with results of other research (Rosenberg and Ekman, 1994) that claims that a person’s subjective report of emotions is coherent with displayed facial expressions. In relation to the topic on how emotions influence human behaviour, including non-verbal behaviour there are different opinions. There is debate regarding the place of emotions relative to behaviour which is well discussed by Baumeister et al. (2007). The authors disagree with an existing statement that the primary function of emotion is to cause behaviour directly. They suggest a view of emotions as a feedback system that operates mainly by means of its influence on cognitive processes which in turn serve as inputs into decision and behaviour regulation processes.
Although a large number of studies on emotion recognition have concentrated on recognising human emotions from facial expressions (e.g., Ekman and Friesen, 1969; Adolphs, 2002; Batty and Taylor, 2003), the facial expressions are not the only source of information that can reveal other people’s emotion. There are many different studies which indicate that the human visual system is capable of detecting, identifying and interpreting biological motion (e.g., Johansson, 1973; Bingham, Schmidt, and Rosenblum, 1995). Human motion can be perceived even from a simple representation of it, such as a point-light display stimulus. That is a method whereby people are videotaped whilst wearing reflective markers positioned strategically about their bodies and perform complex movements in the dark, such as dancing, running or walking (see Figure 1.3). Point-light displays have been used frequently in subsequent studies related to gait perception. In these studies such point-light displays have also been found to contain enough information to convey emotions from
gait, posture and limb movements (Walk and Homan, 1984; Montepare, Goldstein, and Clausen, 1987; Pollick, Paterson, Bruderlin and Sanford, 2001; Atkinson et al., 2004). The results of previous research on the recognition of emotions from human movement are of particular interest to the present study and will therefore be discussed in this chapter.

The issue of recognising affective state from human movement, including whole body movement and limb movement has been explored in many different studies (e.g., Walk and Homan, 1984; Atkinson et al., 2004). A key result of those studies was the finding that there are some emotions that can be determined from human body movements and human gait much quicker, with greater concordance across people, and with less error than other emotions.

![Figure 1.3. Example of visual presentation of Point-light displays in profile view (adapted from Johansson, 1973)](image)

For instance, Walk and Homan (1984) asked participants to label videotapes composed of point-light displays of figures which represented different dances and emotions. The expression of anger was the only emotion that was recognized by participants in the free labelling condition. These results were confirmed by a number of other researchers. Montepare et al. (1987) found that
some emotions were identified on the basis of gait information. For instance, heavy-footedness contributed to the recognition of anger in gait; proud and angry gait was distinguished by greater stride length; and a small amount of arm swing was accompanying sad gait.

Atkinson et al. (2004) provide additional evidence that basic emotions are readily identifiable from body movements, both under normal conditions and in point-light displays. The actors in these studies were asked to portray emotions, with their faces covered, at different levels of exaggeration. The identification of emotions from full-body movements was compared with that from point-light displays. Exaggerated body movements led to an enhancement of accuracy in emotion recognition, especially for point-light displays. Additionally, it was found that by offering a complete depiction of the emotion, in comparison to when offering a point-light display, the recognition of anger and fear was enhanced.

Recognition of affective states in general and negative affective states, such as anger and fear, in particular, is important for humans in social interaction with each other (Mack and Rock, 1998). It is also crucial that people are not only able to perceive what others are doing but also can infer from their gestures and expressions what they might intend to do. As described previously the results of psychological experimental research show that human actions communicate intentions and affective state. The evidence from neuropsychological studies suggests that the brain may be important in mediating human emotions (Damasio, 1994), for instance the negative emotions such as anger (Hariri, Bookheimer, Mazziotta, 2000). The neuropsychological research demonstrates that the healthy animal and human brain seems to react to perceived intentions and feelings quite effortlessly, at a subconscious level. Previous research shows that if someone is being exposed to threat-related behaviour of others, the processing of this behaviour is automatic and rapid and may be mediated by the amygdala (Whalen, Rauch, Etcoff, McInerney, Lee, Jenike, 1998). The amygdala is a brain region located in the medial temporal lobe, and is believed to be involved in the processing of threat in birds, mammals, and reptiles. Activation of human amygdala has been
reported during presentation of angry (e.g., Hariri et al., 2000) and fearful (e.g., Morris, Frith, Perrett, Rowland, Young, Calder, Dolan, 1996) faces, even when stimuli were presented below the level of conscious awareness (Morris, Ohman, Dolan, 1999). Another common finding in the literature is that the results of some functional magnetic resonance imaging (fMRI) experiments which used the manipulation of attention to emotional pictures showed that stimuli which contained fearful faces elicited amygdala activation (it responded more to fearful than neutral face expressions) even when the stimulus was left unattended by participants and they were unaware of it (Vuilleumier, Armony, Driver, and Dolan, 2001; Anderson, Christoff, Panitz, De Rosa, and Gabrieli, 2003). Another study (e.g., Pessoa, McKenna, Gutierrez, and Ungerleider, 2002) also showed that the amygdala can differentiate among facial expressions, but only when the secondary task that the participants were asked to do simultaneously with emotion-discrimination task, was disturbing. To summarize, amygdala activation in response to subtle emotional stimuli such as photographs of facial expressions can be seen as a representation of affective information processing (Davis and Whalen, 2001). These findings also suggest that a high sensitivity in the perception of emotions might facilitate the processing of threat-related stimuli and so increase the sustained attention to this, which is a vital part of social interaction.

The studies which implemented functional neuroimaging provide evidence that distinct cortical areas are involved in the processing of information related to the human body. Cells in the superior temporal lobe seem to respond to images of complex body movements, such as walking, dancing and throwing (Jellem, Baker, Wicker, and Perrett, 2000; Wachsmuth, Oram, and Perrett, 1994). Similarly, a number of fMRI studies have found cortical activations that respond more to biological motion (e.g., walking) compared to non biological motion (Grossman and Blake, 2002). Moreover, various lines of evidence show that the brain of human and non-human primates process a perceived biological movement in such a way that this makes people and animals automatically infer intentions of others (Blakemore and Decety, 2001). The existence of so-called ‘mirror neurons’ (the possible
suggested function of which is to enable an organism to detect certain mental states of observed conspecifics) shows that during the observation of other peoples' movements the same cortical areas are being activated as if the observers themselves were performing this movement (Gallese and Goldman, 1998; Kohler, Keysers, Umlita’, Fogassi, Gallese, Rizzolatti, 2002).

The abovementioned studies provide evidence on how affective state is reflected in non-verbal behavioural cues, such as facial expressions and human motion (e.g., gait, gestures, arm movements, etc.), and how this non-verbal behaviour is recognized by other people. In general, human skills relating to non-verbal communication include the ability to encode or transmit non-verbal cues to other people and the ability to decode or interpret the non-verbal cues transmitted by others (Zuckerman, DeFrank, Hall, Rosenthal, 1978). The ability to decode others’ emotional expressions accurately through non-verbal behaviour has been widely studied in different areas of psychology (Hall and Bernieri, 2001). In regard to the task of recognition the concealed firearm carrying a measure of decoding abilities deems to be important and needs to be examined. Since CCTV operators are monitoring soundless images, it can be assumed that their ability to discriminate between non-verbal cues should be measured by appropriate tests, i.e. the test which does not use any vocal cues. The choice of the suitable test of sensitivity to non-verbal cues will be discussed later, in chapter Two.

Furthermore, as research suggests, the facial and the bodily cues differ in the type of information about affective state which they convey to an observer (Ekman, 1965, Ekman and Friesen, 1967). According to this research, information about the type of emotion can be obtained by perception of the head (i.e. facial expressions), whilst information about the intensity of emotion derives from observation of human body movement or position. It may imply that the dividing of perceived non-verbal information into these two channels can be applied to the task of detecting the bearer of a concealed firearm. Accordingly, the investigation of visual attention distributed between the two main cues, the face and body, of a surveillance target in order to discern their
affective state represents another important topic that is directly related to the
task of recognising the carrying of a concealed firearm.

CCTV cameras typically offer a sub-optimal view of a surveillance target which might systematically influence the judgements of CCTV operators. It can be assumed that recognition of an offender in general and detection of his/her intentions in particular might be dependent on the position and viewpoint of surveillance cameras. This assumption that different viewpoints might influence the recognition of surveillance targets’ actions through decoding of their affective state associated with those actions can be drawn from a number of experimental investigations on the subject of viewpoint dependence in observers’ judgments about the gender, actions and affective state of others represented by point-light displays (e.g., Mather and Murdoch, 1994).

1.3.2. Viewpoint dependence in perception of human gender, actions and affect

Although the effect of the viewpoint on the observers’ perception of a biological motion represented by point-light displays has not received broad attention in the research literature, there are some studies which have systematically investigated the identification of human actions and body postures from different viewpoints. In the following section those studies will be discussed. A number of experimental studies have been conducted to understand the perceptually significant differences between male and female walking patterns (e.g., Mather and Murdoch, 1994; Troje, 2002).

Several studies on the topic of viewpoint dependence in perception of human motion have been conducted in the context of gender classification. Mather and Murdoch (1994) showed that the frontal view, compared to sagittal (i.e., side view) or other views, provides an observer with enough information about a walking person’s gender. These experiments were inspired by previous work on the topic of gender discrimination. In the earlier work (e.g., Kozlowski and Cutting, 1977; Cutting, Proffitt, Kozlowski, 1978) it was found that perceptual differentiation between male and female walking patterns were
mediated by cues associated with the human torso. According to Mather and Murdoch (1994) the frontal views ought to be more effective and informative about the cues related to movement of the human torso than the side view. Similar to these findings, the results of the study of Troje (2002) showed the advantage of frontal view in the task of gender recognition from biological movement.

Moreover, there is some evidence that different angle of viewing human posture gives rise to different perception of human action. The study of Daems and Verfaillie (1999) focuses on the representations underlying action and pose identification. Based on the previous neurophysiological (e.g., Perrett, Harries, Bevan, Thomas, Benson, Mistlin, Chitty, Hietanen, and Ortega, 1989; Wachsmuth et al., 1994) as well as psychological (e.g., Verfaillie, 1993) evidence they assumed that depending on the orientation and position of the acting body relative to the observer, different instances of a particular action or pose may produce very different retinal images, but will have to receive identical labels nevertheless. In their study the identification of human actions and body postures viewed from different viewpoints was examined in four long-term priming experiments with static pictures of a human model. In a number of sequential experiments the participants had to name or describe the pictures, and to decide whether the pictures showed a possible or impossible body pose. Having seen the same action or pose in a different orientation did not reliably facilitate identification performance later on. Also, there was no priming for poses that are impossible to perform with a human body, not even when an identical same-view prime was used. Daems and Verfaillie (1999) concluded that these findings suggest that the stored representations that mediate the identification of human actions and postures are viewpoint specific.

In parallel with the viewpoint dependence research in recognition of gender and human actions, the question about the role of viewpoint dependence in human affect recognition arises. Coulson (2004) suggested that in contrast with other channels of non-verbal communication, body posture involves an important three-dimensional presence which offers different percepts depending on the observer’s location. Changes in viewing angle may
therefore result in occlusion of one body part by another. The previously
discussed study performed by Daems and Verfaillie (1999) showed evidence
which suggests that the same posture viewed from different angles does not
give rise to the same percept, as the prior presentation of a posture does not
prime its later recognition from a different viewpoint. With respect to affect
recognition from body postures, Coulson (2004) showed that there was a
general tendency for frontal views to lead to more consensual attributions of
affective states to presented postures, although this pattern did not apply for
every presented affective state \(i.e.,\) for fear and sadness). According to
Coulson (2004) these results suggest that “attributing emotion to a body
posture is a great deal easier when the person adopting the posture is facing
the perceiver. Such an orientation, while not necessarily ideal for perceiving the
three dimensional relationships between body segments, may nonetheless
enhance recognition due to its interpersonal significance. … When we are
confronted, rapid and accurate decoding of emotional state offers important
information that can be used to guide behaviour.” (p. 135-136).
The studies mentioned in the current and the previous sections show that the
recognition of human actions/gender and recognition of human affect from
those actions might be dependent on the viewpoint. Consequently, the
inclusion of viewpoint as an additional variable in the current research was
deemed relevant, assuming that the observation and recognition of actions and
affective states of surveillance targets might also depend on the view the
surveillance targets appear to an observer on the CCTV imagery.

1.4. The present research

The present research was part of the larger MEDUSA project the aim of which
was to reduce gun crime by developing of a new automated, machine learning
system that would be able to interpret human behavior and to identify threat
related situations and individuals from CCTV footage in real time. Although, as
it was mentioned earlier (see section 1.1), there are many obstacles that affect
the performance of human CCTV operator in a negative way, and the fact that the firearms are rarely spotted by CCTV operators in the course of proactive surveillance (Darker et al., 2008), the firearms are being carried in public places with considerable frequency. Thus, there is a great necessity in multidisciplinary research which will investigate the possibilities to improve the proactive detection of crimes involving firearms via CCTV. One of such possibilities is a software which will allow to identify firearm related threats and to alert the human operator.

As part of the MEDUSA project, the development of such software was intended to be achieved, among other things, through identifying overt and covert perceptual and cognitive cues which human operators use to identify the behavior related to gun crime, and through combining human and machine-derived cues throughout development of the intelligent software. In this respect, the overt perceptual and cognitive cues refers to the cues which can be verbally described and therefore communicated easily between individuals; the covert cues are subconscious cues, which due to their nature cannot be easy described and thus they are not communicated between individuals. In the research of the current thesis main attention was driven to the covert cues, such as change in non-verbal affective behavior that may accompany carrying a firearm. Detection of people engaged in criminal activities (e.g., carriers of illegal firearms) through careful observation of their affective state and non-verbal behaviour have been already largely recognized to be important in the work of police officers (Burns, 2006; Pinizzotto, Davis, Miller, 2006; Johnson, 2007; Remsberg, 2007). However, so far, insufficient applied research has been done with respect to the work of CCTV operators, which may include detection and responding to incidents with illegal use of firearms. Previously discussed review of published work showed the ability of a weapon to automatically produce an affective, negative response, and discussed the issue of recognition of affect from non-verbal behavoural cues such as facial expressions and human motion (e.g., gestures, gait etc.). Expressive behaviour seems to provide people with knowledge about the potential reactions and the intentions of others, and possibly also with information about future behaviour.
Taken together, it suggests that the CCTV operator’s ability to predict an unlawful act accompanied with using a firearm may be linked to the ability to recognize the affective and in particular the negative affective state of others through non-verbal expressions.

The main aim of the present research is to investigate whether the ability of a human operator to identify the behaviours that indicate that someone is carrying a concealed firearm might be based on the analysis of body language related cues (e.g., facial expressions, change in gait), and particularly on an understanding of negative affective state derived from these cues. Since it is believed that observers would respond to subtle cues which individuals inherently produce whilst carrying a concealed firearm, it is also desirable to identify the possible cues which accompany the carrying of concealed firearm, and to understand the techniques and abilities people might use in order to differentiate between a bearer of concealed firearm and a bearer of innocuous object. Another assumption of the current research is that the ability to recognize the carrier of concealed firearm in the CCTV footage might be affected by other factors, such as the skills in decoding affective state of others and the viewpoint of observation the surveillance targets.

To summarise, in order to investigate the possible link between the recognition of human affect and the detection of a concealed firearm carrier it is necessary to undertake research on the following issues: the ability to recognise concealed firearm carrying amongst experts (i.e. CCTV operators) and lay-people; the relationship between sensitivity to non-verbal cues and the ability to detect the carrying of a concealed firearm; the recognition of concealed firearm carrying in conscious and sub-conscious ways, and the influence of viewpoint dependence on this task; the identification of perceived visual cues associated with concealed firearm carrying based on the recognition of the affective state of surveillance targets. In order to give an empirical basis to this research a number of experimental studies have been performed and will be described in the following chapters. The structure of this thesis is presented in the next section.
1.5. Thesis structure

The thesis presents a number of experimental studies which have been carried out on the basis of the theoretical evidence derived from previous scientific research discussed in the current chapter. Chapter 2 is mainly concerned with the methodology used in the research. It describes the idealized CCTV footage of people carrying firearms that was made for the purpose of the investigation. The clips were generated for use in the experimental studies and as part of the MEDUSA project. Additionally, the chapter describes the assessment tool that was used in order to measure the sensitivity to non-verbal cues amongst observers, their ability to recognise the affective states expressed by others, and a number of other questionnaires used in this research.

Chapter 3 describes an experiment which was performed to investigate the effect of carrying a firearm on the affective state of the surveillance targets. The research question is based upon evidence that suggests that the act of committing a crime and carrying an illegal firearm is associated with certain affective responses in offenders. A self-report measure of affective state was used in order to examine whether carrying a gun was accompanied by certain affective states that can be reported by the gun carrier themselves. The results of the study are discussed with regard to the current investigation.

Chapter 4 investigates the abilities of CCTV operators and lay people to identify an individual who is carrying a concealed firearm through CCTV imagery. By assuming that this identification might be based on body language analysis, in this study an assessment tool was used which measures sensitivity to non-verbal cues in order to investigate the emotion decoding ability of participants. In particular, the chapter presents a study which investigates the performance of CCTV operators and lay people in the task of detecting a concealed gun carrier (i.e. Gun Carrier Detection task, GCD task) from idealized CCTV footage, and the effect of expertise in the field of CCTV surveillance on this performance. This experiment is further discussed in order to demonstrate the relationship between the affect recognition abilities of observers and the conscious recognition of the bearer of a firearm from CCTV
imagery. Performance on the gun carrier detection task is assessed within a signal detection framework.

Chapter 5 describes a study carried out in order to find evidence for this relationship \((i.e.\) affect recognition – recognition of concealed firearm carrying) by approaching the issue from a different angle. The ability to discriminate between a person carrying a gun and a person carrying an innocuous object is inferred from performance on a task involving the recognition of the affective state of surveillance targets \((i.e.,\) Affect Detection task, AD task). In other words, the task of observers is to estimate the affective state of surveillance targets in the mock CCTV footage without knowing about the presence of firearms. Any potential effect of the observers' awareness of the presence of a firearm on their decision making process is therefore minimized. Additionally, a test of body language decoding ability was performed to examine whether the body language reading skills of observers were related to their ability to detect a concealed firearm via affect detection.

Chapter 6 discusses a study carried out in order to examine further the task of affect recognition in relation to the detection of a concealed firearm carrier. This issue is investigated by using footage of the same surveillance targets taken from a different viewpoint \((i.e.\) street-level perspective). Another aim of this study was to investigate which visual cues the observers use to perform this task, and in particular which visual cues are associated with the presence of a firearm. This information about visual cues is obtained from questionnaires regarding inspection strategies implemented whilst performing the task.

In Chapter 7 these issues are further investigated and discussed with regard to visual cues associated with the recognition of the affective states of the surveillance targets and level of performance on the concealed firearm detection task. In addition to previous data obtained from the questionnaires, information about visual cues used in this task is now obtained using eye-movement registration and analysis.

Chapter 8 presents the overall discussion and conclusions of this research, limitations and indications for further work.
Examples of all the questionnaires used in this research are given in the Appendices.
CHAPTER 2: METHODOLOGY

2.1. Introduction

This chapter describes the methodology used in the current research. As mentioned in the Introduction, the thesis presents a number of experimental studies which have been carried out in order to investigate whether the recognition of offenders with a concealed firearm by a human operator might be based on the recognition of affective state on the basis of silent, non-verbal behaviour that is accessible from CCTV images. The empirical evidence is needed to support the assertion that human observers might apprehend and respond to subtle cues which individuals might inherently produce whilst carrying a concealed firearm, which in turn might be reflected in the body language of those carrying firearms.

With respect to the affective response to carrying a firearm it can be assumed that experience with carrying a firearm might be an influential factor that diminishes the degree to which a bearer of a firearm would emotionally respond to its presence. The more experienced users of guns or professionals (e.g. police officers) can be expected to be less affected by carrying a firearm than novice users. As mentioned earlier, nowadays there is a large number of inexperienced, ‘amateur’ criminals who possess, carry and use firearms (Lockhart et al., 2007). Taking this fact into consideration, in the current research the influence of firearm carrying on affective state was investigated by using people who were not experienced in using a gun.

An important aspect which needs to be taken into account in the research on the estimation of others’ or one’s own affective state is that individuals can be predisposed to a high level of aggression by certain traits (Anderson and Bushman, 2002). An initial assessment of the general trait aggression of participants was therefore necessary in order to control for its
possible effect on the self-rating of negative affect as well as on peer-ratings of participants who served as observers.

Another assumption which needs to be examined is that the ability to recognize the carrier of a concealed firearm in the CCTV footage might be affected by other factors, such as the skills in decoding an affective state of others or the viewpoint of observation the surveillance targets. In the following sections the details of the questionnaires and tests that were applied in the current research and detailed description of other materials are given.

The purpose of the first part of the current chapter is to present information obtained through a number of visits to CCTV centres in several cities in the UK. Together with the literature reviewed in Chapter 1 this helped set the background to the current research. The second part of the chapter will describe the generation of mock CCTV footage of persons whilst walking with either a concealed firearm or an innocuous object matched to the firearm, and to describe the Signal Detection Theory that was used to analyse the performance of observers on the task of detecting a concealed firearm carrier. As part of the MEDUSA project an experimental set up was designed to create the video clips which were used in the experiments relating to the MEDUSA project. Regarding the present research, the filming sessions also served to investigate the effect of carrying a firearm by inexperienced people on their affective state as compared with carrying an innocuous object, and to investigate the possibility of detecting a carrier of a concealed firearm through observation of their affective state.

The third part of the chapter introduces the questionnaires that served as a measure of the estimation of an affective response of people whilst carrying a firearm or an innocuous object (matched to the firearm). The following part will describe the test selected in order to measure sensitivity to non-verbal cues, and will discuss the rationale behind the selection of the test. Finally, the chapter presents the method (i.e. eye movement registration) used to collect the information about visual cues relevant to the tasks of affect and concealed firearm carrying recognition.
2.2. Survey of CCTV control rooms

Several CCTV control centres were first visited to gain an understanding of how CCTV operators work, how they are selected and trained. The centres selected were: Charnwood Borough Council, Nottingham CCTV control room and a CCTV control room in Liverpool. Several repeated visits were made to these control rooms. These were selected as follows, on the basis of their links, or lack of links, to gun crime. Charnwood is the local CCTV centre which deals with Loughborough, a largely rural town with little history of actual gun crime. Nottingham has a long standing reputation as a city with a major gun crime problem which in recent years has seen a decrease in prevalence. In earlier years rival drug gangs in certain areas of Nottingham had made it popularly known as the gun capital of England. In 2002 it had the eighth highest number of gun crimes in England and Wales. In the four years between 2002 and 2006 Nottinghamshire Police seized over 350 firearms, recovered some 3,500 rounds of ammunition and detained more than 600 gun crime related offenders. Liverpool is a large city which also has considerable gun crime, which like Nottingham has seen a decrease in recent years. According to Home Office data, 12% of recorded gun crime (excluding air weapon offences) occurred in Greater Manchester, 4% in Merseyside and 2% in Nottinghamshire (Hales et al., 2006). In 2004-2005 56% of all crime in England and Wales took place in the area of London, the West Midlands, Greater Manchester and Nottinghamshire. Home Office data also show that the number of crimes involving firearms recorded by police nationally rose by 11% from 10,023 in 2001 to 11,084 in 2005. In 2005, Merseyside had recorded a 183% rise in the number of persons found guilty of possessing or distributing prohibited weapons or ammunition between 1997 and 2005 (Brogan, 2007).

Taken together then the three centres provided a wide breadth of CCTV operator opinions concerning their work and how this relates to potentially identifying someone carrying a gun. In particular, methods of spotting suspicious human behaviour, including carrying and using illegal firearms, were investigated. Also operators’ opinions were canvassed with regards to the
current gun crime situation in their city and the potential for detecting it via CCTV.

From these visits it became clear that only a small percentage of incidents involved firearms and were spotted by CCTV operators proactively (i.e. when they identified suspicious individuals from CCTV images and informed the relevant security staff or the police about this). A situation occurs regularly when operators need to respond to police requests to focus a camera on a certain place, so instead of being pro-active the operators often act reactively. In general, the operators confirmed that it is very difficult to detect a person with a firearm from the large amount of live CCTV images that they need to oversee, and also due to their other responsibilities (e.g., carrying out administrative tasks, maintaining CCTV storage or communication with others).

According to the CCTV operators in Liverpool, they are not looking at all of the available 60 CCTV screens all the time, but just at some critical moments of the day and at certain critical places in the city. For example, at 9am they pay extensive attention to post offices in the city and at 4pm the operators reported that they look at parking places more often than elsewhere. Usually, when a gun related crime has been detected proactively then this has occurred when the CCTV operators were monitoring an event for a different reason than just a suspicion of gun crime. Nevertheless, according to operators, once an unsafe situation on the street has been recognized, then they would try to estimate the chance that a shooting could take place. In particular in cities where there has been a large gun crime issue (for example, Nottingham) the CCTV operators were fully aware of the possible presence and usage of firearms. Moreover, the CCTV operators in Liverpool control room reported that there was a detection of a shooting incident by operators at least once a week.

According to one of the senior operators in the Nottingham CCTV control room certain qualities or ‘innate abilities’ of operators play an important role in spotting mal-intent in general and especially in recognizing an individual with a firearm; making them able to perform well on this kind of task irrespective of any training. With regard to training issues it was noted that many of CCTV operators are former police officers and therefore have previous experience
with numerous types of offences, often including offences where a firearm was involved. In accordance with a competence-based training approach (Diffley and Wallace, 1998) all new recruits underwent practical assessments to ensure they possess the necessary competences (e.g., ability to detect conspicuous details from CCTV imagery). Subsequently, the new selected operators received training which typically involved different techniques such as attending a 3-4 day training course, after which the new operators also had ‘on the job’ training. This followed the guidance on training CCTV operators, which emphasises a combination of different instructional techniques which is known to be the most appropriate for CCTV operators (Diffley and Wallace, 1998).

In addition, the visits to control rooms also gathered real-life CCTV footage for the MEDUSA project which had been collected by CCTV operators. In particular, the Charnwood CCTV control room provided footage that showed young people openly carrying firearms on the street and incidents where people were threatened with guns. The review of real-life footage obtained over several years from police forces as well as local authorities nationally showed people, mainly young men, walking along the street carrying firearms either openly or concealed. Therefore this appeared to be an ideal point to start in firearm detection as it might capture the casual carrying of firearms before their use. This real-life CCTV footage of potential gun crime situations from CCTV control rooms and police forces demonstrated the need to generate an idealized CCTV footage for the purpose of experimentation. The recording method that was used is described in the following section.

2.3. Creating the mock CCTV footage of people carrying firearms

Because of the restricted time period for the collection of the requested amount of CCTV footage and the fact that the real-life footage of gun crime is rare and difficult to get hold of, it was decided to generate staged CCTV video clips of
walking people carrying real firearms and innocuous, matched objects to be used in the experiments related to the MEDUSA project (e.g., Darker, Gale, Blechko, 2008; Blechko, Darker, and Gale, 2008; 2009). This was done in order to create CCTV footage for use in experimental scenarios that would mimic the live, real-time detection of individuals carrying firearms either overtly or covertly on the person. The research attention on detecting firearms carried by a walking person originated from the assumption that this would be within the capability of CCTV, whereas detecting firearms, for example, in vehicles might be better suited to other technology. The below mentioned location, cameras, recording method, wardrobe, firearms and matched objects were chosen in order to create the mock CCTV footage of 12 male surveillance targets.

**Location**
The videotaping took place in a room with restricted access at Loughborough University, under the supervision of two level-five certificated firearms experts (one from Royal Armouries, and one from LGC Forensics, both sited in Leeds, UK). The area filmed comprised a 10.8 m walkway marked on a grey floor, and a white-painted breezeblock wall which served as a back-drop. The room was lit by three fluorescent tube style ceiling lights. These pre-existing room lights were supplemented by two spotlights which front-lit the scene. The reason for choosing the location inside of the building refers to the requirements of the MEDUSA project. This would enable machine-vision algorithms, which would be based on the findings of the research on human visual inspection strategies in the detection of firearms, to be developed better. Furthermore, the choice of this location seemed also to be reasonable as carrying of a firearm can always occur inside of buildings (e.g., shops, hotels, airport, bus or train stations etc.) along with carrying on the streets. It also permitted the creation of a set-up which could be well controlled in terms of experimental design. Expert advice was obtained from firearms professionals about the set up. This choice of location was a good imitation of a real-life situation since the carrying of any firearms including air weapons is strictly prohibited in the UK (Violent Crime
Reduction Act, 2006), which makes the making of such experimental video footage in public places extremely difficult.

**Cameras**

As discussed previously (see chapter One) there are a number of factors that are known to influence the performance of CCTV operators within a security control room. Along with human-related factors (e.g., communication, management, training and experience of CCTV operators), factors related directly to the CCTV technology may either enhance or reduce the performance of operators (Gill et al., 2005; Cohen, Gatusso, and MacLennan-Brown, 2009; Keval and Sasse, 2010). In particular regarding crime reduction, some characteristics of CCTV technology can have a great impact on the effectiveness of CCTV against crime. Examples of such characteristics are the temporal quality of CCTV imagery (i.e., temporal compression, or the interval between the images that are stored) and its spatial quality (i.e., spatial compression, or the number of pixels used to make up the image). Furthermore, the relative size that the target person appears on screen which depends on the type of observation, lighting design, and the field of view in terms of optimal camera placement, are important as it may also affect significantly the observation and identification of an individual from the CCTV imagery (Cohen *et al*., 2009).

In the current research the participants were filmed using two cameras. Camera 1, a PAL CCTV video camera, was positioned to obtain a wide perspective of the walkway from a three-quarter angle (the visual angle of 53.6°), at 7.45 m from the walkway and from a height of 2.79 m (see Figure 2.1). There was a sub-optimal lighting created in order to mimic a live feed. This camera produced CCTV footage at 50 frames per second with minimal compression, without pause, rewind or zoom-in. According to the CCTV Operational Requirements Manual 2009 (Cohen, *et al*., 2009) a PAL camera that captures 25 and higher frames per second give the appearance of smoothly flowing motion, and is more than adequate for most situations. Camera 1 was intended to mimic the set-up of a real CCTV camera, the
passive CCTV view, before any detection is made and before CCTV operators start zooming in. Since the making of this video footage was a part of the MEDUSA project, it should be mentioned that the MEDUSA automatic software system was intended to make an initial detection for unattended cameras. However, there are some problems with this view such as the fine details (e.g., the face of firearm carriers) that would of necessity not be too visible. This means that the camera view can reveal information about the utility of such macro-scale cues such as gait and gross movements, and to much lesser degree about reading emotions from the face. According to the CCTV Operational Requirements Manual 2009 (Cohen et al., 2009) a figure in the image should occupy at least 5% of the screen height and the scene portrayed should not be unduly cluttered to enable an observer to monitor the direction and speed of movement of people across a wide area, and at least 10% of the screen to be able to detect a person. The images that were produced by the camera in the current study correspondence to these requirements. In the video images each surveillance target subtended approximately 4° visual angle and the firearms and bottles subtended between 0.1 and 0.5 visual angle, depending on type, at a viewing distance of 450mm.

Figure 2.1. Schematic representation of the filming set-up
Camera 2 was positioned at 3.1 m from the start point of the walkway and at a height of 1.6 m. \textit{(i.e.} at head height\textit{)}. This camera was used to obtain the video footage of the same surveillance targets but made from different viewpoint \textit{(i.e.} street-level perspective\textit{)}. This additional street-level camera view was used in order to obtain additional information about the use of the face of firearm bearers as a cue in the affect recognition task since the faces of the targets were more visible in this view compared to camera 1. The CCTV Operational Requirements Manual 2009 (Cohen \textit{et al.}, 2009) recommends placing the camera at head height when the suspect identification is the main priority, as ceiling mounted cameras may not be able to provide a full view of the individual’s face. The detailed description of the obtained footage is given in section 2.6.3. and Chapter 7 (section 7.2.1.2.).

\textit{Recording}

The digital signal from camera 1 was conveyed from the Canopus digital video converter (Grass Valley, version ADVC 110, manufacturer Canopus Co., Ltd, 2004-2006) to the computer via Firewire connector (6 pin; IEEE 1394) and was sampled at 25 fps. The resultant video stream was captured straight to hard drive in a compressed DV (with compression level 10:1) “avi” format. This process was performed using 'DV Rack’ software by Serious Magic (2005, version 1.2.1490.0) on a computer with Microsoft XP (version 2002, Service Pack 2, manufacturer Viglen Ltd.; Inter ® Xeon [TM] CPU 3.20 GHz; 2.00 GB of RAM).

Another video camera (camera 2) was used to produce the video footage of the frontal view of surveillance targets. This was a Sony Handycam Optical 20x / 800x digital zoom, Model no. DCR-HC 22 E.

\textit{Surveillance targets}

Twelve male students or staff members from Loughborough University volunteered to be filmed as surveillance targets. They were between 23 and 35 years of age \textit{(M = 26.8 years, SD = 3.4)}, one of them was left-handed. Ten of the twelve participants had previous, but not significant experience with
firearms (including air weapons). All the participants were naive to the purpose of the research and were paid for their participation.

According to the recent statistics obtained in England and Wales and presented by Metropolitan Police Authority (Hales et al., 2006) gun crime is mainly committed by young men. Due to this fact it was decided to use only male surveillance targets. The present scenario (i.e. lone person walking) was used to concentrate specifically on the potential detection of firearms on the person (e.g. as observed in review of real life footage) and to leave out any reactions and influences of bystanders.

Wardrobe

The surveillance targets wore different types of clothing which they could choose before the first filming session. The “clothing collection” available for this experiment consisted of leather and parka jackets, and ‘hoodies’ of different colours (see Figure 2.2).

![Figure 2.2. Examples of clothing outfit used in the experiment](image-url)
This type of outfit was chosen in consultation with firearm experts and on the basis of a review of real-life footage of firearms incidents as undertaken in Darker et al. (2007).

**Firearms and matched objects**

As suggested by the gun crime experts who provided their advice on the set-up of the experiment, many models of handgun can be easily concealed on a person and are often used by criminals who wish to carry a handgun for illegal purposes and by those who bear firearms for self-protection. Three firearms were used in this study (see Figure 2.3). They were genuine firearms, provided by the Royal Armouries in Leeds: a small revolver, a self-loading pistol (Glock), and a sawn-off shotgun. The firearms were not loaded with ammunition and there was no ammunition on site.

Each firearm was matched with an innocuous every-day object that approximated its size, weight and colour. The following three types of firearms have been used: .32 Caliber Revolver (170 mm, 0.53 kg), Glock (200 mm, 0.94 kg), and Sawn-off Shotgun (390 mm, 1.86 kg). The firearms were selected in consultation with a firearms expert to represent those most used in street gun crime.

*Figure 2.3. Presented firearms: A – .32 Caliber Revolver; B – Glock; C – Sawn-off Shotgun*
To match the firearms in terms of size and weight soft-drink bottles (including contents) were used (see Figure 2.4). Three sizes of bottles have been used: \( \frac{1}{2} \) -litre (230 mm, 0.5 kg), 1-litre (275 mm, 1 kg), and 2-litre (340 mm, 2 kg).

![Figure 2.4. Presented matched objects – bottles of different sizes and weights](image)

**Special precautions**

Special precautions were taken on account of safety issues; for example, two, level-five certificated firearms experts, one from the Royal Armouries and one from LGC Forensics, Leeds, were present at all times. Participants were recruited from a small pool of known, trusted people, without widely advertising the experiment. Further, the local constabulary and university security had been informed about the time and the location of the experiment. There was also controlled access to the filming area.

2.4. Measuring the performance on the task of recognizing a concealed firearm carrier by observers: Signal Detection Theory

The performance of observers on the task of recognising a concealed firearm carrier was analysed using Signal Detection Theory (SDT). Signal Detection Theory is widely applied in psychology, medicine, and other related fields. It is a standard model for predicting detection and discrimination in a search paradigm. SDT is used to analyse data derived from experiments where the task is to categorize ambiguous stimuli which can be generated either by a
known process (*i.e.*, the signal) or by chance (*i.e.*, the noise) (Abdi, 2007). In the present research two conditions were created: signal present (*i.e.*, the concealed firearm is featured in the clip) and signal absent (*i.e.*, the concealed bottle is featured in the clip). According to SDT, the noise condition is a condition when no signal is present. SDT also assumes that the signal is added to the noise (Abdi, 2007). In the present experiment the presence of an object featured in the clip that is not a firearm (*i.e.*, an innocuous object matched in size and weight to firearm) reflects the signal absent condition, or noise only condition. Accordingly, the presence of an object that is a firearm reflects the signal present condition which is equivalent to a signal plus noise condition.

An attempt was made to match the signal present and signal absent conditions on the basis of the size and weight of the object carried. It was assumed that the signal would reflect visual cues other than those related to the presence of a concealed innocuous object of the same size and weight as a firearm. These visual cues might be the behavioural consequences of the affective state evoked by carrying a firearm. An important parameter for the SDT model is the value $d_a$. It is a standardised (z-score based) measure of the distance between the means of the noise and noise plus signal frequency distributions, over a range of confidence levels relating to confidence that the signal is present or absent. It represents a parameter that is numerically equal to $d'$ when the two distributions are of equal variance (RSCORE 5.3.2; Harvey, 2001). Since the parameter $d_a$ describes the relationship between the noise and signal distribution, it measures how readily the signal can be detected (Wickens, 2002) and can thus be used as a measure of discriminability between signal and noise (*i.e.*, a measure of sensitivity to the signal). Therefore, in order to measure the sensitivity to firearms in the concealed firearm carrier detection task the parameter $d_a$ was used.
2.5. Measurement of affective state

As mentioned in the chapter One, the literature suggests the existence of a ‘weapon effect’ which is known to be characterised by the ability of a weapon to automatically prime aggression. As a part of the present research, an attempt has been made to find out whether carrying a firearm indeed elicits emotional (i.e., affective) and, in particular, negative response in a bearer. Therefore two relatively new, well validated questionnaires have been used to measure the general trait of aggression of participants and their affective state during filming sessions: the Aggression Questionnaire (Buss and Perry, 1992) and the Multiple Affect Adjective Checklist – Revised (MAACL-R; Lubin and Zuckerman, 1999) respectively.

2.5.1. Aggression Questionnaire

The Aggression Questionnaire (AQ; Buss and Perry, 1992) is a self-report trait measure of aggression which meets current psychometric standards and has reasonable correlations with peer-ratings of aggression (the total score of correlation was .31 with the highest correlation score of .45 for Physical Aggression scale; Buss and Perry, 1992). This was used to assess the general trait of aggression of participants in Experiment 1 (see Chapter 3) in order to control whether the emotional state reported by participants is primarily due to the effects of experimental manipulations and not to their general level of aggressiveness. This test is a validated rating scale measuring four components of aggressiveness: physical (Physical Aggression scale, 9 items), verbal (Verbal Aggression scale, 5 items), anger (Anger scale, 7 items), and hostility (Hostility scale, 8 items). The 29 items on the AQ concern self-reports of behaviour and feelings, with each item scored using a 5-point Likert scale. The total score for Aggression is the sum of the factor scores. The examples of statements are: “If I have to resort to violence to protect my rights, I will” (physical aggression); “I tell my friends openly when I disagree with them” (verbal aggression); “Some of my friends think I am a hothead” (anger); “When
people are especially nice to me, I wonder what they want” (hostility). For all questions see Appendix 1.

2.5.2. Multiple Affect Adjective Check List- Revised (MAACL-R; state and trait form)

The Multiple Affect Adjective Check List- Revised (MAACL-R; Lubin and Zuckerman, 1999) is a well known self-report measure of affective states. Therefore this assessment instrument (i.e., the State form of MAACL-R questionnaire) was chosen in order to measure the emotional state of participants in the present research while carrying a gun compared with carrying a matched innocuous object of similar weight and size. Besides that, the MAACL-R Trait questionnaire was used in order to control whether or not the level of surveillance targets’ trait aggression predisposed toward their self-ratings of affect during the experiment. Furthermore, the Trait form of the MAACL-R questionnaire was also used to assess the general trait of aggression of participants who served as observers in the present experiments in order to control whether the affective state reported by these participants had any effect on their perception of affective states in the surveillance targets, bearing in mind that rating someone else’s affective state can be vulnerable to the rater’s own traits. According to Epkins (1994), one of the important factors that can bias the human observer’s ratings of other people affective states is the tendency to rate others high or low on the traits on which the observers rate themselves as high or low (i.e., the so-called ‘Projection error in peer-ratings’).

Because an earlier version of this test (MAACL) has been criticized due to high intercorrelations among the three negative affect subscales (i.e., anxiety, depression, and hostility) the new revised version of this test (MAACL-R) contains two additional composite scales: Dysphoria, which is a construct created by combining the raw scores on Anxiety, Hostility, and Depression subscales, whilst a positive affect is a construct represented by the scores on Positive Affect and Sensation Seeking subscales. Each subscale score equalled the number of mood-relevant adjectives checked. As it appears that
the overlap among the negative affect scales has been reduced, the MAACL-R test is now being widely used in psychological research. The MAACL-R consists of 132-items (all are adjectives) which represent five, abovementioned subscales: anxiety, depression, hostility, positive affect and sensation seeking. The estimation of the negative effect or Dysphoria can be done by combining the raw scores for anxiety, hostility, and depression. There were 10, 12, 15, 21, and 12 mood adjectives of the anxiety, depression, and hostility subscales, respectively. Each subscale score equalled the number of mood-relevant adjectives checked. Examples of adjectives include afraid, fearful (Anxiety subscale), lonely, rejected (Depression subscale), angry, annoyed (Hostility subscale), affectionate, friendly (Positive Affect subscale), and active, adventurous (Sensation Seeking subscale). Approximately one-half of the items are distracters. For an example of the MAACL-R questionnaire see Appendix 2 (State form) and 3 (Trait form), and detailed instructions see Chapter 4 and 5.

2.6. Measurement of sensitivity to non-verbal cues of human observers: Profile of Nonverbal Sensitivity (PONS) test

A large body of research has investigated the ability to judge another’s emotional expressions (Hall and Bernieri, 2001). There exist several well-established tools for measuring non-verbal sensitivity. Most such measures of non-verbal sensitivity assess accuracy in the recognition of emotions as expressed by others, known as decoding skills.

With respect to the current research the choice of the proper measurement to assess accuracy in the recognition of emotions as expressed by others, was based on advice of the expert in the fields of non-verbal cues to deception (personal communication, 22.11.2006, Prof. Aldert Vrij; University of Portsmouth). There were also other considerations. There was a need to find a test which, in statistical terms, would have a reasonable reliability. It should also
discriminate itself from other tests in one important aspect, namely it should not use vocal cues. Vocal cues are a part of non-verbal communication, but they are not important in terms of monitoring CCTV images; a CCTV-operator’s main responsibility is monitoring and analyzing the soundless, live images on screens for, among other things, subtle changes in human non-verbal behaviour.

For the purpose of the current research it was decided to use the well-known test of the affective decoding abilities, Profile of Nonverbal Sensitivity (PONS) test (Rosenthal, Hall, DiMatteo, Rogers, and Archer, 1979). The overall aim of the PONS test is to identify which people are more or less vulnerable to non-verbal cues in social interaction with other people. Those with lower PONS scores are less sensitive to such cues and can be referred to as poor ‘decoders’. They might have a greater tendency to miss or misjudge the cues exhibited by other people. The test measures, in particular, the ability to decode facial expressions, body movements, and tone of voice, both separately and in combination. Tone of voice was not of interest here and only silent clips are supposed to be used in the present research.

The reliability of the test has been assessed in terms of stability over time and internal consistency, which was .86 for the full-length PONS test. Retest reliability of the full version of test over periods of 10 days to 10 weeks is .69 (median over 6 samples). The short forms of this test have an internal consistency reliability of .40, this is weaker compared with full-length test due to fewer items (40 vs. 220). However, according to Hall (2001) most non-verbal decoding tests seem to lack strong internal consistency; this is not just a characteristic of the PONS test.

With respect to the validity of this short test, Hall (2001) assumed that at present it is not possible to argue that one decoding ability test is more valid than another. In addition, the PONS was subject to an extensive programme of validational research. Hall (2001) also gives brief descriptions of several other tests which made it possible to compare different approaches to measuring decoding abilities, and which also gave a good basis for justifying the selection of the PONS test for the present study. Compared to other non-verbal decoding
ability tests which are concerned with only emotion judgments (e.g., Diagnostic Analysis of Nonverbal Accuracy, DANVA, Nowicki and Duke, 1994; Japanese and Caucasian Brief Affect Recognition Test, JACBART, Matsumoto, LeRoux, Wilson-Cohn, Raroque, Kooken, Ekman, Yrizarry, Loewinger, Uchida, Yee, Amo, and Goh, 2000) the PONS test taps a wide variety of domains and emotions, as well as social scripts, status roles, deception, relationship cues. Additionally, it addresses different non-verbal channels within the same test which all contribute to a good ecological validity. An important advantage of this test is the possibility to investigate separately the sensitivity to facial expressions and to bodily cues, assuming that the face can often be barely discernible via CCTV (albeit it is possible for an operator to zoom in on face), and body movements are most easily seen.

Specifically the shortened version of the test was administered in the current research and only those portions which address ability to read body language from either the face or body, on the basis of visual cues alone. This version of the test comprised 40 items: 20 face-only and 20 body-only items from the full version of PONS test. The items are 3.5-second video fragments of a 24-year-old woman acting in different naturalistic, emotional situations, from which the verbal content is excluded. The role player enacted each situation expressing either positive or negative affect in a dominant or submissive manner. The face and body video-fragments of the PONS test each represent two dimensions, a positive versus negative dimension and a dominant versus submissive dimension. Examples of descriptions characterizing the situations presented in the video: 'Admire nature' (positive/dominant); 'Helping customer' (positive / submissive); 'Expressing jealous anger' (negative / dominant); and 'Asking forgiveness' (negative / submissive). After viewing the video fragment the observer is required to make a choice between two descriptions for each video presentation, as quickly and as accurately as possible. The observer is instructed to pick the statement that best described the situation in which the emotion is portrayed. An example of this two-alternative forced choice is a choice between ‘expressing motherly love’ and ‘asking forgiveness’. In the present study the Face and Body PONS was digitized, therefore the videos as
well as the descriptions of enacted scenes were presented on the computer screen. The inter-item interval was not limited, but the participant was asked to choose the answer as quickly and as accurately as possible. Three measures of non-verbal sensitivity were derived from this test: proportion correct across all 40 items; proportion correct across the 20, face only items; and proportion correct across the 20, body only items. The PONS face score and the PONS body score were used separately in the analysis in order to obtain information on decoding skills.

2.7. Detection the visual cues related to concealed firearm carrying

2.7.1. Cue-detection questionnaires. Questionnaire design

The recognition of positive or negative affect in others usually involves an identification and rating of visual characteristics of gait and overall movement patterns such as, for example, stride lengths, arm swing or speed of movement (Montepare, Koff, Zaitchik, and Albert, 1999). There was therefore a need to understand how exactly carrying a firearm influences the appearance of the carrier to the observer in terms of cues related to carrier's emotional state. The simple way to investigate this was by presenting to the observers a questionnaire that can collect data on which cues convey information needed to identify affective state in a firearm bearer and a bearer of an innocuous object. The aim of the questionnaire was thus to find out how observers’ estimation of mood of gun and no gun carriers is related to their observation of different properties of surveillance targets’ appearance and behaviour. In particular, the questionnaire served to examine how characteristics of targets’ movement pattern (e.g., strides’ length, degree of arms’ swinging) are related to:

1 - estimated by observers affective state of surveillance targets;
2 - type of task (carrying a firearm versus carrying a innocuous object).
The design of the questionnaires was based on previous work that assessed the visual characteristics of gait. For instance, Montepare et al., (1987) asked observers to identify emotions from point-light displays (for the description of the point-light displays see Chapter 1, section 1.3.1.), and to make ratings of four gait characteristics: stride lengths; arm swing; heavy-footedness, and posture. They found that angry gait was distinguished by heavy-footedness and by greater stride length. Montepare et al. (1999) found significant effects of emotion type across different movement ratings (e.g., jerky, loose, fast, hard, contracted). The results showed, for example, that angry clips were rated as the most jerky whereas sad and neutral clips were rated as the most smooth. In the present study an assessment of movement speed was also adopted in the questionnaire as it has been indicated that velocity change seems to influence the perception of affect from the arm movements of another (Pollick et al., 2001; Paterson, Pollick, and Sanford, 2001). In their study it was found that movement speed modulates the intensity of angry affect perceived in arm movements, whilst sad and neutral movements were affected by changes in velocity in such a way that the faster sad or neutral movement were categorized as angry. In addition, Meijer (1989) showed that each distinct emotion category (e.g., joy, fear and anger) was predicted by a unique combination of movement features. The arm movement, for example, was a predictor of fear, whereas a lower degree of fear was associated with a higher degree of arm movements.

Based upon the abovementioned studies the cue-detection questionnaire in the current research consisted of two parts. In the first part the questions were asked about the movement pattern of the surveillance targets. Movement pattern was specified by eight characteristics (i.e., cues): length of stride, degree of arms’ swinging, degree of heavy footedness (related to surveillance targets’ gait); degree of fluidity, degree of rigidity, degree of speed, and degree of exaggeration (related to the surveillance targets’ overall movements).

In the second part questions were asked about how the observers gained an impression of the affective state of surveillance targets by looking at
different cues (i.e., gait, posture, facial expressions, position of the arms, and speed of walking). This could provide information about to which degree the use of those cues underlies the observers' estimation of affect in surveillance targets. The latter was done by applying a five-point Likert response scale: (e.g., I looked at their gait) not at all; rarely; sometimes; often; all the time. For all the questions of the cue-detection questionnaire see Appendix 4. The questionnaire provided a first indication as to how the identification of a surveillance targets’ affective state differs in terms of the visual cues used by observers while performing this task. As mentioned earlier, carrying a firearm seems to increase negative affective state (e.g., hostility which can be related to anger). The degree of ‘arms’ swinging,’ speed of walking, and length of stride could be associated with level of negative affect, and therefore with carrying a concealed firearm. Although the questionnaires provide an initial indication concerning the answers to the research questions, they do not provide a complete picture. In order to fully understand the strategies which observers are using during this task their visual behaviour should be explored using more reliable, empirical method, such as eye-tracking. By doing this the information about which parts of the image are relevant for performing this task could be inferred from observers’ eye-movement pattern, without relying solely on the observers’ consciously reported strategies, as is the case with the questionnaire method. Therefore, an eye-tracking method was employed in this research.

2.7.2. Eye tracking method

An effective way to expand the relevant data was to involve the measurement of eye movements of observers. Eye movements has been seen as a tool of great value for psychologists as it can provide researchers with a rich, dynamic data source concerning the temporal dynamics and psychological processes that led up to the response (Richardson and Spivey, 2004). Furthermore, the eye tracking technique is known to be able to clarify whether eye-movements are related to attention to emotionally arousing information presented in images
(Nahm, Perret, Amaral, Albright, 1997; Green, Williams, Davidson, 2003). Applying this particular method can, therefore, also provide information about whether the perception of a particular affective state in the CCTV video clip is associated with a certain eye-movement pattern. The eye-movements pattern was examined by estimating the average time the observers spent on looking at different areas surrounding specific bodily features (e.g., face) which can be seen as a measure of amount of attention to head or body of surveillance targets. It provides information about which cues people are using to detect surveillance targets’ affective state.

**Apparatus**

A Tobii X50 Stand-alone eye-tracker with ClearView 2.6.0 software was used to collect the eye movement data. The Tobii system comprises two screens connected to the computer allowing participants to view the stimulus on one screen (i.e., stimulus computer screen) while the other screen (the eye-tracking computer screen) can be used by the experimenter to set up the test and watch the eye-tracking quality during the session. For the set up of the experiment see Figure 2.5.

*Figure 2.5. Set-up of the eye-tracking experiment*
Stimulus presentation was done on the 17” computer display with a maximum resolution of 1280 x 1024 pixels. For the purpose of the study the video clips were presented with low resolution of 800 x 600 pixels. Participants sat approximately 60 cm away from this display screen which provides the stimulus. The Tobii Eye Tracker has a maximum accuracy of 0.5 degrees and a maximum sampling rate of 50 Hz (Tobii User Manual, 2006).

The data collection on the task of affect detection was automated using another computer, i.e., a questionnaire computer (Toshiba Tecra, Toshiba Ltd.; Microsoft Windows XP version 2002; Genuine Intel ® Centrino DuoTM T2600 processor 2.16 GHz; 994 Mhz motherboard) and controlled by a program developed in-house.

2.7.3. Mock CCTV footage used in the experiment with eye movements registration

The mock CCTV footage comprised multiple video clips of 11 actors (i.e., surveillance targets), each filmed individually and whilst walking and carrying either a concealed firearm or a concealed innocuous object matched to the firearm for approximate weight and size (i.e., a bottle). Each video clip appeared in the middle of the computer screen in the form of a rectangle, the size of which was 23 cm in height and 18 cm in width. At a viewing distance of 60 cm this corresponds to 24.6 and 19.3 degrees of visual angle in width and height respectively. Each participant was shown a surveillance target walking towards them in the video clip. The range of the mean height of the targets varied between 9.3 cm (9.95 degrees of visual angle) [at the start point of walking path] and 12.7 cm (13.6 degrees of visual angle) [at the end point of walking path]. The range of the mean width of the targets at the broadest points (i.e., across two forearms) varied between 3.4 cm (3.6 degrees of visual angle) [at the start point of walking path] and 4.5 cm (4.8 degrees of visual angle) [at the end point of walking path]. The duration of each video clip was 4 seconds. On account of the on-screen size of the targets, accurate registration of eye
movements in each area of interest was only possible for the two last seconds of each video clip.

2.8. Conclusion

The current chapter provides a description of the methodology applied in the present research. Details of participants, equipment, questionnaires and test used in the experimental studies are given. In particular, the chapter describes the mock CCTV footage that was generated of persons (i.e., surveillance targets) whilst concealing either a firearm or an innocuous object matched to the firearm. The footage was used in the experimental studies related to the MEDUSA project and the current research. Besides, it reveals the choice of the measurement of sensitivity to non-verbal cues, the self-measure of the affective state and the questionnaires applied in order to collect the data about the visual cues related to the task of recognition of affect in humans carrying a firearm and the task of detection of firearm carrier. In addition, the chapter also demonstrates the rationale behind the preference for selected tests.

This thesis so far has drawn upon the existing evidence from the research literature to highlight the possible relationship between perception of and distinguishing between affective states in humans and the identification of a bearer of concealed firearm through CCTV imagery. In the following chapters the research which contains a number of experimental studies will be described and the results discussed.
CHAPTER 3: EXPERIMENT ONE

3.1. Introduction

In the introduction to the thesis the evidence from applied research was reviewed showing that the act of committing a crime is associated with certain emotions. A number of studies (e.g., Canter and Ioannou, 2004; Hales et al., 2006) revealed that certain affective processes amongst offenders were found to be associated with the act of committing a crime or carrying a firearm. These emotions appear to vary with the type of crime; exhilaration was especially associated with robbery and property crimes, whilst anger was prominent for violent crimes including murder (Canter and Ioannou, 2004). Moreover, the existence of the ‘weapon effect’ has been demonstrated which is known to be characterised by the ability of a weapon to automatically prime aggression. The literature reviewed previously shows (see Chapter 1) that with respect to firearms offences, seeing a firearm is sufficient to automatically prime aggressive thoughts (Anderson et al., 1998) and to promote aggressive behaviours (Berkowitz, LePage, 1967; Klinesmith et al., 2006). Thus, it might be inferred that carrying an illegal weapon as a strong sensory stimulus may be transformed to an inner representation (i.e., a sensation) in the offender. The literature review presented earlier provides a theoretical basis for conducting experimental studies in order to examine whether carrying a firearm compared to carrying an innocuous object elicits an affective response in a carrier and whether this response can be recognized by other human observers. The study presented in this chapter describes an attempt to find out whether carrying a gun elicits emotion (i.e., affective response), and in particular, negative emotional responses in a gun carrier. This study forms the first, basic study of the thesis but was designed and performed as an embedded part of the MEDUSA project.
The research question which was investigated was: *Is carrying a gun accompanied by the experience of certain affective, and in particular, negative states which can be reported by a gun carrier themselves?* Another purpose of the experiment was to obtain mock CCTV footage of people carrying firearms. This part of the experiment was described earlier, in the Methodology Chapter (see Chapter 2, section 2.3). As mentioned before, an experimental set up was designed to generate the mock CCTV footage of individuals who acted as surveillance targets whilst concealing either a firearm, an innocuous object matched to the firearm in terms of weight and size, or no object. For the purpose of the current study the filming sessions served to investigate the effect of carrying a firearm by inexperienced people on their affective state compared to carrying an innocuous object. In order to measure participants’ affective state during the filming sessions the Multiple Affect Adjective Checklist – Revised State (MAACL-R State; Lubin and Zuckerman, 1999) was used. Since certain traits are known to predispose individuals to a high level of aggression (Anderson and Bushman, 2002) it seemed to be necessary to measure the trait aggression of participants. Therefore, two questionnaires were used: the Aggression Questionnaire (Buss and Perry, 1992) and the MAACL-R Trait (Lubin and Zuckerman, 1999).

3.2. Method

3.2.1. Participants

Twelve male students and staff members from Loughborough University volunteered to participate in the study. They were between 23 and 35 years of age ($M = 26.8$ years, $SD = 3.4$), one of them was left-handed. Ten of the twelve participants had previous, but not significant experience with firearms (including air weapons). All the participants were naive to the purpose of the study and were paid for their participation.
3.2.2. Materials

A detailed description of the materials (i.e., location of filming, video recording and cameras, wardrobe of the surveillance targets, types of firearms and matched objects) used in this study is given in Chapter 2, section 2.3.

Emotional State Measures

The following questionnaires were completed by each participant: Aggression Questionnaire (Buss and Perry, 1992) and the Multiple Affect Adjective Check List- Revised State Questionnaire State and Trait (MAACL-R; Lubin and Zuckerman, 1999). For the description of these measures see Chapter 2, section 2.5., and Appendix 1 (Aggression Questionnaire), Appendix 2 and 3 (MAACL-R State and Trait forms). The Aggression Questionnaire and MAACL-R Trait questionnaire was used in order to assess trait aggression in the surveillance targets. The State version of MAACL-R questionnaire was used to measure the affective state of participants while carrying a gun compared with carrying a matched innocuous object. For this purpose the Hostility scale of both questionnaires was used. When participants were given the MAACL-R State form they were instructed to check those mood adjectives that described how they felt whilst we were filming them, immediately after filming session. The questionnaire took 3–5 minutes to complete.

3.2.3. Procedure

Some 792 clips across the 12 participants were made (see Table 3.1 and 3.2) using camera 1 (for schematic representation of the filming set-up see Chapter 2, figure 2.1). The participants were assigned to filming sessions according to the following principle: all the participants with an odd number (Group A) performed the Neutral (walking whilst carrying no items), the Gun (walking whilst carrying a firearm), and the Innocuous Object (walking whilst carrying a bottle) sessions, whereas the participants with an even number (Group B) performed the Neutral, Innocuous Object, and Gun sessions. Each participant
was filmed separately; the order of gun and bottle filming was counter-balanced across participants: participants in Group A carried the guns first; the participants in group B carried the bottles first. Thus, the effect of carrying a gun compared to a bottle was counterbalanced across participants.

<table>
<thead>
<tr>
<th></th>
<th>Neutral</th>
<th>Innocuous Object (Bottle)</th>
<th>Gun (Gun)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unconcealed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking towards the camera</td>
<td></td>
<td>108</td>
<td>36</td>
</tr>
<tr>
<td>Walking away from the camera</td>
<td>108 x towards</td>
<td>108</td>
<td>36</td>
</tr>
<tr>
<td><strong>Concealed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking towards the camera</td>
<td></td>
<td>108</td>
<td>36</td>
</tr>
<tr>
<td>Walking away from the camera</td>
<td>108 x away</td>
<td>108</td>
<td>36</td>
</tr>
</tbody>
</table>

*Table 3.1. Overall number of the clips that were filmed across all participants.*

Every filming session consisted of a specified number of ‘walking trials’ (see Table 3.2). The first session, ‘Neutral’, which served as a training session and was the same for all the surveillance targets, consisted of 18 walking trials (9 trials towards the camera and 9 trials away the camera). The second and the third filming sessions consisted of 12 walking trials.
<table>
<thead>
<tr>
<th></th>
<th>Neutral</th>
<th>Innocuous Object (Bottle)</th>
<th>Gun (Gun)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unconcealed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking towards the camera</td>
<td>3 x ½ litre</td>
<td>1 x .32 Revolver</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 x 1 litre</td>
<td>1 x Glock</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 x 2 litre</td>
<td>1 x Sawn-off shot</td>
<td>gun</td>
</tr>
<tr>
<td>Walking away from the camera</td>
<td>3 x ½ litre</td>
<td>1 x .32 Revolver</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 x 1 litre</td>
<td>1 x Glock</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 x 2 litre</td>
<td>1 x Sawn-off shot</td>
<td>gun</td>
</tr>
<tr>
<td></td>
<td>9 x towards</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Concealed</strong></td>
<td>9 x away</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking towards the camera</td>
<td>3 x ½ litre</td>
<td>1 x .32 Revolver</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 x 1 litre</td>
<td>1 x Glock</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 x 2 litre</td>
<td>1 x Sawn-off shot</td>
<td>gun</td>
</tr>
<tr>
<td>Walking away from the camera</td>
<td>3 x ½ litre</td>
<td>1 x .32 Revolver</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 x 1 litre</td>
<td>1 x Glock</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 x 2 litre</td>
<td>1 x Sawn-off shot</td>
<td>gun</td>
</tr>
</tbody>
</table>

*Table 3.2. Overall number of clips made with every kind of carrying object (i.e., three types of bottles and three types of firearms) in each session, for one participant.*
There was a fourth filming session (‘Innocuous Object’) involving only bottles which was the same for all participants. Additionally, every filming session that involved carrying a gun or bottle was divided into ‘concealed’ and ‘unconcealed’ walking trials: sometimes the participants were instructed to walk six times with the item (firearm or bottle) in the hand and visible to the cameras (‘unconcealed’); and another six times the participants were asked to hide the item under the clothing before walking (‘concealed’). In the fourth session there were 12 concealed and 12 unconcealed walking trials with innocuous object. In total 66 video clips were made of each participant. Each recording lasted approximately 10 seconds. The fourth session was necessary to obtain the correct number of clips for a different statistical design.

During all of the ‘Gun’ filming sessions two representatives from the Royal Armouries and LGC Forensics, first gave short explanations concerning the firearm and advised the participants about the way of concealing a firearm from view. During the ‘Innocuous Object’ ‘concealed’ filming session the surveillance targets were free to decide how to hide the bottle themselves.

Running order for participants in both groups, A and B, of participants

1. Buss and Perry Aggression questionnaire and MAACL-R Trait questionnaire
Two days before the experiment all the participants received the Aggression questionnaire and the MAACL-R Trait questionnaire that they were asked to fill in and bring along on the day of the experiment.

2. Filming session one: Neutral clips
The experiment was conducted over two consecutive days. Participants came to the room individually. Before the participants were given the first instructions, they had to read and sign the Informed Consent form. However, the purpose of the experiment was withheld to avoid influencing behaviour. Each participant (i.e., surveillance target) was asked then to change into the ‘walking outfit’, and the participant’s number, age and ethnicity were noted. The surveillance target was then instructed to walk towards and away
from the camera whilst being videotaped, as naturally as possible, at their own pace.

3. **MAACL-R state**

After the first session the participant was asked to fill in the MAACL-R State questionnaire.

4. **Filming session two: group A (Gun / Bottle) and group B (Bottle / Gun)**

Once this was completed, the surveillance target from group A (i.e., Gun- Bottle sequence) was instructed to do the same again, except this time he was asked to walk one time towards and away from the camera holding a small revolver visible to the camera in the hand (filming session ‘Gun’, ‘unconcealed’), and one time - with a small revolver, covered by clothing (filming session ‘Gun’, ‘concealed’). This procedure was repeated with two more different kinds of firearm (i.e., Glock and sawn-off shotgun). In total the surveillance target had to walk with a firearm six times towards the camera and six times away from the camera.

The surveillance targets from group B (i.e., Bottle-Gun sequence) were instructed to do the same as the surveillance targets from group A, but in this session were required to carry bottles of different sizes.

5. **MAACL-R state**

Subsequent to this filming session the MAACL-R State questionnaire had to be filled in by the participant.

6. **Filming session three: group B (Bottle /Gun) and group A (Gun / Bottle)**

After finishing the questionnaire, the procedure was repeated, but this time instead of three different kinds of firearms the surveillance target from group A was asked to hold (or to hide) bottles of three different sizes (1/2 litre, 1 litre and 2 litres) whilst walking (filming session ‘Innocuous Object’, ‘unconcealed’ and ‘concealed’). The surveillance target from group B was given the same instructions but instead of bottles he had to carry guns of three different types.
7. **MAACL-R state**
At the end of this filming session the participant had to fill in the MAACL-R State questionnaire.

8. **Filming session four: remainder of bottle clips.**
When the surveillance target had finished the questionnaire, the surveillance target was asked to walk towards and away from the camera holding a bottle (of three different sizes) for another 24 times (Filming session ‘Innocuous Object’, ‘Concealed’ and ‘Unconcealed’). Both the walking trials with concealed and unconcealed bottle were repeated 12 times.

9. **Debriefing – true purpose revealed**
Once the last filming session had been done, each surveillance target was instructed to change out of the walking outfit and was thanked for their participation in the experiment. After that the participant received a financial inconvenience allowance, and was asked to put their address, ID number and signature on a financial form. Additionally, the participant received the debriefing sheet with the true purpose of the experiment revealed, the questionnaire about the participant’s previous experience with firearms, and directions to the University Counselling Service in case they experienced any stress during or after the experiment.

### 3.3. Results

Statistical analysis was performed adopting an $\alpha$-level of .05. After controlling the data for normality of distribution the Wilcoxon Signed Ranks Test was performed in order to determine whether the level of affect assessed by the MAACL-R State questionnaire varied as a result of gun carrying. The results of this test suggest that the level of Dysphoria varied significantly across two conditions, *i.e.* when carrying a firearm and when carrying an innocuous object ($z = 2.68$, N-ties = 9, $p = 0.007$). When the participants’ scores on the
component subscales of Dysphoria were analyzed separately, it showed similar effect of a carrying subject on all three subscales ($z = 2.55$, N-ties = 8, $p = 0.011$, for Anxiety; $z = 2.06$, N-ties = 5, $p = 0.039$, for Depression; $z = 2.07$, N-ties = 5, $p = 0.038$, for Hostility). No significant effect of the carrying object on the PASS and subscales Positive Affect and Sensation Seeking scores were found. Table 3.3. represents the distribution of MAACL-R scores on different scales (i.e., the 25th and 75th percentiles, medians, maximum and minimum scores) across two conditions, Gun and Innocuous Object.

The scores on Dysphoria, Anxiety, Depression and Hostility were significantly higher in the Gun condition than in the Innocuous Object condition, which means that after the filming sessions when participants had to carry a gun they reported feeling more anxious, depressed and hostile than after the filming sessions when they had to carry a bottle. There were no significant differences in their experiences of positive affect or sensation seeking between the Gun and Innocuous Object sessions.

<table>
<thead>
<tr>
<th></th>
<th>Anxiety (Innocuous Object)</th>
<th>Anxiety (Gun)</th>
<th>Depression (Innocuous Object)</th>
<th>Depression (Gun)</th>
<th>Hostility (Innocuous Object)</th>
<th>Hostility (Gun)</th>
<th>Dysphoria (Innocuous Object)</th>
<th>Dysphoria (Gun)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.67</td>
<td>2.25</td>
<td>0.25</td>
<td>1.08</td>
<td>0.67</td>
<td>1.25</td>
<td>1.58</td>
<td>4.58</td>
</tr>
<tr>
<td>Median</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Minimum</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Percentiles 25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>Percentiles 75</td>
<td>1.75</td>
<td>4.5</td>
<td>0.75</td>
<td>2</td>
<td>1.75</td>
<td>2.75</td>
<td>3</td>
<td>9.75</td>
</tr>
</tbody>
</table>

*Table 3.3.* MAACL-R scores on dysphoria, anxiety, depression and hostility scales for two conditions: Gun and Innocuous Object.

**Correlations among the scores on the Aggression and MAACL-R state questionnaires**

The Spearman correlation test was carried out with the participants’ scores on the Hostility scale of the Aggression Questionnaire and their scores on Hostility scale of MAACL-R Trait questionnaire. This was not significant ($r_s = 0.314$, $n = 12$, $p = 0.320$, two-tailed). The same analysis was applied to examine
relationship between the participants’ scores on the Hostility scale of MAACL-R Trait questionnaire and their scores on the Hostility scale of MAACL-R State. This analysis did not show a significant relationship ($r_s = 0.523$, $n = 12$, $p = 0.081$). This result showed that the level of their trait aggression did not predispose toward their self-ratings of negative affect during the experiment.

### 3.4. Discussion

In accordance with results of previous studies, the results of this experiment showed that when participants were carrying a gun, they reported feeling significantly more anxious, depressed and hostile than during the filming sessions when they had to carry an innocuous object. The trait aggression of participants, measured prior to the experiment, showed no correlation with their affective state during the experiment. Furthermore, there was no relation found between the participants’ general level of hostility and the reported level of hostility during experimental trials when they had to carry a gun. It can therefore be concluded that the experience of holding a gun during the experiment had contributed to the self-reported negative affect experienced by participants whilst performing the task.

### 3.5. Conclusion

Previous research showed that carrying a firearm by offenders was sufficient to be associated with a certain emotional response by an offender (Cusson, 1993; Canter and Ioannou, 2004; Hales et al., 2006) and to increase the accessibility of aggressive thoughts by ordinary people (Anderson et al., 1998). The study described in this chapter demonstrated that carrying a gun compared with carrying a neutral, harmless object can elicit differences in the measured affective state of a carrier. In terms of statistical significance, this effect was
present related to for a negative mood state. The level of trait aggression of participants seemed to have no relationship with the affective state during the experiment. The obtained correlation between the scores on the Hostility scales of both questionnaires showed no relationship. This implies that the general affective state and attitudes do not need to be seen as a necessary precondition for the increase of negative mood resulting from carrying the mood-eliciting object. In other words, the results of this experiment gave an indication that carrying a gun was associated with an increase in negative mood in a carrier; which is in accordance with the hypothesised predictions.
CHAPTER 4: EXPERIMENT TWO

4.1. Introduction

The overall aim of the present chapter is to investigate the abilities of CCTV operators compared to lay people in identifying an individual who is carrying a concealed firearm. It was assumed that this identification by CCTV operators might be based on body language analysis and particularly on an understanding of the emotional state of the target person as derived from subtle behavioural cues. To ascertain the potential influence of expertise and training the performance of both CCTV operators and lay people were examined on this task.

The current research on the topic of surveillance through CCTV (e.g., Donald, 1999) provides evidence that training is likely to enhance individuals’ abilities, but those who have natural potential for this kind of work appear likely to maintain the performance gap even after training, compared with people who do not have this natural predisposition. Donald (1999) evaluated within a CCTV control room the utility of the assessment exercises measuring the core competencies of CCTV operators (i.e., speed of information processing; ability to effectively monitor and deal with a situation; accuracy in rapidly identifying and responding to representation of people's behaviour). Donald demonstrated that the results of training evaluation assessments conducted in his study indicated that there are underlying sets of competencies within people that make them naturally more suitable for surveillance and detection work in a control room. Thus, it appears that in order to make the training of CCTV operators successful, it is necessary to identify and to take into account individual differences in specific skills and natural abilities, and accordingly to be able to decide to what extent they need to be trained. One of the important individual characteristics which underlies the skills of the surveillance task is the ability of a CCTV operator to analyse and understand the intentions of others.
through observation of non-verbal behaviour. It is possible that the ability of experts, i.e. CCTV operators, to read body language and non-verbal behaviour is highly developed due to their training and work experience.

As outlined in the introduction chapter, a number of studies have revealed that certain affective processes amongst offenders have been found to be associated with the act of committing a crime (e.g., Canter and Ioannou, 2004). The literature reviewed previously (see chapter One) shows a ‘weapons effect’ that means that seeing a firearm is sufficient in itself to increase the accessibility of aggressive thoughts (e.g., Anderson et al., 1998). Study One (see Chapter 3) showed that this can be demonstrated empirically simply by asking people to walk with a real firearm and then measuring their affective state afterwards. The existing psychological research and the literature regarding the experiences of police with criminals indicates that there exists a belief that it is possible to predict the intents and actions of offenders by reading subtle non-verbal cues in the appearance of a person (Burns, 2006; Pinizzotto et al., 2006). It has been suggested that these cues can be meaningless for a novice but are significant for a trained person (Ekman, 2003). However, other researchers (e.g., Troscianko et al., 2004) demonstrated that the detection rates for incidents and antisocial behaviour perceived via CCTV footage was very high for both experienced CCTV operators and lay people. Such results suggest that training and experience do not necessarily make a difference in detecting mal-intent via CCTV. Thus, it might be inferred that the recognition of affective non-verbal cues related to carrying a firearm is related to the innate decoding abilities of observers.

The next section describes the study which was performed in order to investigate the potential influence of expertise by comparing the ability of CCTV operators with those of lay people in the task of detection of individuals carrying a concealed firearm based upon skills in reading affective, non-verbal cues. Additionally, in order to examine the relationship between individual differences of CCTV operators compared to lay people in their sensitivity to non-verbal cues and the performance on the task of detection of firearm carrier a test of affect decoding abilities (i.e. PONS test) was performed.
4.2. Experiment Two

4.2.1. Introduction

The experiment\(^1\) reported here focuses upon determining whether visual recognition can support the detection of firearms, via CCTV, and whether this ability is more highly developed in CCTV operators due to their previous training and experience in the surveillance task. Consequently, the decoding ability of participants was measured and the relationship between their sensitivity to non-verbal cues and their performance on the gun detection task was investigated.

The abilities of CCTV operators and lay people to detect concealed firearms in video clips derived from mock CCTV footage (for the full description of the footage see Chapter 2, section 2.3. were assessed within a signal detection framework. It was hypothesised that CCTV operators would demonstrate greater sensitivity in the detection of concealed firearms than lay people. If this hypothesis could be confirmed, it might suggest that the skills involved in this task can be acquired through training or experience.

Additionally, an influence of the affective state of the surveillance target on the judgements of CCTV operators and lay people regarding the detection of a concealed firearm was investigated. It was hypothesised that the number of times a surveillance target was deemed to be carrying a firearm would correlate with a metric derived from their affective state (\(i.e.,\) negative affect in terms of anxiety, depression, and hostility, and positive affect in terms of positive affect and sensation seeking). If this relationship could be found, then this would suggest that the affective state of the surveillance target is manifested in a visually apparent form and that the related visual cues are used by an observer in order to detect a firearm. Finally, the sensitivity to the firearm was assessed relative to a standard measure of the abilities of CCTV operators

\(^1\) The paper that described this study ‘Skills in Detecting Gun Carrying from CCTV’ has been published by Blechko, A., Darker, I.T., Gale, A.G. in the Proceedings of ICCST 2008 42nd Annual IEEE International Carnahan Conference on Security Technology
and lay people to read the affective content of body language and facial expressions (i.e. non-verbal affective cues). Would those with greater sensitivity to non-verbal cues exhibit greater sensitivity to the concealed firearm?

4.2.2. Method

4.2.2.1. Participants

Sixteen people (two groups, CCTV operators and lay people) volunteered to participate in this study. Eight participants were CCTV operators (age: $M = 37$ years, $SD = 10$) and 8 were lay people (age: $M = 47$ years, $SD = 12$). All the CCTV operators were employed at CCTV control rooms in the UK (years of experience as a CCTV operator: $M = 5$ years, $SD = 3$ years). The lay people had no training or experience in any sector of the security industry. All participants were naïve to the purpose of the study.

4.2.2.2. Materials

Apparatus
Stimulus presentation and data collection were automated using a computer (Toshiba Tecra, Toshiba Ltd.; Microsoft Windows XP version 2002, service pack 2; Genuine Intel® Centrino DuoTM T2600 processor 2.16 GHz; 994 Mhz motherboard; 1.00 GB of RAM) and controlled by a program developed in-house.

The generation of mock CCTV footage and the measurement of sensitivity to a concealed firearm
The mock CCTV footage was described in the previous chapter (see Chapter 2, section 2.3). It comprised multiple video clips of 12 actors (surveillance targets), each filmed individually and whilst walking and carrying either a firearm, an innocuous object matched to the firearm for approximate weight and size (a bottle), or no additional object. For the purpose of the present analysis, only
video clips in which the objects (firearms and bottles) were carried concealed on the person (see Figure 4.1) were used.

The levels of negative affect (anxiety, depression, hostility) and positive affect (positive affect and sensation seeking) were assessed for each surveillance target, both after carrying a firearm and after carrying the innocuous object. Therefore the state version of the Multiple Adjective Affect Checklist – Revised (MAACL-R, Lubin and Zuckerman, 1999) was used. For the detailed description of this test see Chapter 2, section 2.5.2, Appendix 2. This assessment was achieved within an experiment design that was counterbalanced for order of exposure to the firearm and innocuous object.

Figure 4.1. Examples of still images of the idealised footage (A – walking with a concealed firearm; B – walking with a concealed two-litre bottle); (adapted from Blechko et al. (2008))

The order of clips was randomised with respect to the type of object carried (firearm, bottle, none). In the present study two conditions were examined: Gun (i.e., the firearm is featured in the clip, concealed on the surveillance target) and Innocuous Object (the bottle is featured in the clip, concealed on the surveillance target). The measure of sensitivity to concealed firearm was $da$ which is appropriate when the two distributions are of unequal-variance, and is numerically equal to $d'$ when the two distributions are of equal variance (Harvey, 2001). For further explanation of this aspect of the measurements in terms of SDT see Chapter 2, section 2.4.
The measurement of sensitivity to non-verbal cues

Prior to participating in the firearm detection part of the experiment, each participant undertook a standard test of their ability to read body language: the PONS test (Rosenthal et al., 1979). For the detailed description of the test see Chapter 2, section 2.6.

Design

Of interest were the influences of the expertise of the observer, the ability of the observer to read body language and the emotional state of the surveillance target on the decision regarding whether or not a surveillance target was carrying a concealed firearm. Participants were therefore assigned to one of two groups on the basis of their expertise in CCTV surveillance: CCTV operator or lay person.

4.2.2.3. Procedure

CCTV operator participants were recruited by contacting CCTV control room managers and through speaking about the study at CCTV practitioner meetings. Lay person participants needed to be matched to the recruited CCTV operators in terms of age and gender and were recruited by advertising within Loughborough University, from opportunity samples within the locality of Loughborough, and by contacting organisations within Loughborough University.

The experiment consisted of two sessions. Participants were administered the Face and Body PONS test in the first part of the experiment, and the firearm detection test in the second part of experiment. Two parts were run together as one experiment. Each participant took part once. They were first administered the PONS test which took 5 minutes to complete. They then undertook the Gun Carrier Detection test which took approximately 90 minutes (excluding a short break of 5 minutes) to complete.
Session 1 – the PONS test:
The aim of the first session was to assess the participants’ ability to read body language. Body language reading ability was indexed in terms of ability to decode silent, non-verbal, behavioural cues. The PONS test was automated on a laptop computer. In order to familiarize the participants with the procedure the experiment included a practice session of eight items. The practice session was followed by the 40 items of the main test, given in a randomised order.

Session 2 - the Gun Carrier Detection test:
The firearm detection test was presented on the same laptop computer as the PONS test. Participants were required to watch video clips derived from the mock CCTV footage with the aim of measuring their sensitivity to the concealed firearm. It was preceded by a practice session, whereby the participants viewed eight practice video clips. They then viewed 504 experimental video clips: 72 clips of people carrying concealed firearms; 216 clips of people carrying concealed bottles; 216 clips of people carrying no additional object. The video clips were presented in a pseudo-randomised order which spread the signal detection conditions evenly across the time-course of the experiment. It was necessary to have a large number of clips in order to give the study a desired statistical power. The ratio of clips containing no object, and concealed gun or bottle was worked out based on 80 % of power and 95 % confidence of detecting cues associated with a firearm carrying.

After each video clip in the GCD task, each participant had to indicate whether or not the person in the clip was carrying a firearm. They indicated their confidence on a six point scale which was presented on the computer screen: “Definitely no”; “Probably no”; “At a guess, no”; “At a guess, yes”; “Probably yes”; and “Definitely yes”.

71
4.2.3. Results

Statistical analyses were conducted to explore the firearm detection sensitivity of CCTV operators and lay people using an α-level of .05. Potential relationships between the self-reported affective state scores of surveillance targets and the abilities of observers to recognise when those individuals were carrying a firearm were assessed. Finally, potential relationships between an observer’s sensitivity to non-verbal cues and their ability to detect a firearm carrier were assessed.

Expertise and performance on the Gun Carrier Detection task

In order to calculate sensitivity to firearm carrying within a signal detection framework, the frequency of responses across the six response categories in the Gun and Innocuous Object conditions were used. Performance was indexed in terms of the Receiver Operator Characteristic (ROC) parameter $d_a$ which was obtained by means of a maximum likelihood algorithm and based on a Gaussian distribution (RSCORE 5.3.2; Harvey, 2001).

The results of the current study indicates that there was a tendency for both groups of observers, CCTV operators and lay people, to mistake a concealed innocuous object for a concealed firearm as their performance (i.e. sensitivity to concealed firearm) was below chance (see Figure 4.2). However, there was a trend for CCTV operators to have higher sensitivity than lay people in the detection of concealed firearms (See Figure 4.2). Furthermore, sensitivity to the concealed firearm between CCTV operators ($Mdn = -0.31$) and lay people ($Mdn = -0.44$) did not differ significantly (Mann-Whitney $U$ test: $U = 16.00$, $p = .105$, $r = -0.42$). Consequently, CCTV operators and lay people are considered together for the rest of the analysis.
Correlations between the affective states of surveillance targets and the frequency with which they were deemed to be carrying a concealed firearm.

Spearman rank order correlations were performed between the number of times a surveillance target was deemed to be carrying a firearm and the level of affect experienced by that surveillance target, separately for each unique combination of affect scale (anxiety, hostility, depression, positive affect, and sensation seeking) and signal detection condition (Gun and Innocuous Object).

For data pooled across CCTV operators and Lay people, the Spearman rank order correlation tests showed that there was a significant correlation between the Anxiety scores of surveillance targets in the Innocuous Object condition and the number of times that a given surveillance target was deemed to be a firearm carrier ($rs = 0.61, n = 12, p = .034$, two-tailed). In the Gun condition the same test performed between the Anxiety scores of surveillance targets and the number of times that a given surveillance target was deemed to be a firearm carrier (again, for data pooled across CCTV operators and Lay people), showed no significant correlation, but the result did approach significance ($rs = 0.5, n = 12, p = .098$, two-tailed). No other correlations approached significance.
Spearman rank order correlations were also conducted between the number of times a surveillance target was deemed to be carrying a firearm and the size of the influence of carrying a firearm on the surveillance target’s affect level (\textit{i.e.}, a difference between the level of affect experienced by a surveillance target in the Gun and Innocuous Object conditions), separately for each affect scale (anxiety, hostility, depression, positive affect, and sensation seeking). This analysis revealed that the size of the influence of carrying a firearm on the surveillance target’s anxiety level exhibited a significant, positive correlation with the number of times that that individual was deemed to be a firearm carrier \((rs = 0.59, n = 12, p = .042, \text{two-tailed})\) (see Figure 4.3). However, no significant correlations were found on the basis of depression, hostility, positive affect, or sensation seeking.

\textit{Figure 4.2.} Sensitivity \((d_a)\) to a concealed firearm amongst CCTV operators and lay people (adapted from Blechko \textit{et al.} (2008))
Figure 4.3. Correlation between the size of the influence of carrying a firearm on the surveillance target’s anxiety level and the number of times that the surveillance target was deemed to be carrying a concealed firearm (adapted from Blechko et al. (2008))

Correlations between the affective states of surveillance target and sensitivity to that target carrying a concealed firearm.

Spearman rank order correlations were performed between observer’s sensitivity to a surveillance target carrying a firearm and the size of the influence of carrying a firearm on the surveillance target’s level of affect (i.e., the difference between the level of affect experienced by a surveillance target in the Gun and Innocuous Object conditions). It was performed separately for each affect scale (anxiety, hostility, depression, positive affect, and sensation seeking).

The size of the influence of carrying a firearm on the surveillance target’s level of anxiety did not influence the sensitivity of an observer to the concealed firearm when carried by that individual. Rather, the size of the influence of
carrying a firearm on the surveillance target's sensation seeking level demonstrated a significant, negative correlation with the sensitivity of an observer to the concealed firearm when carried by that individual ($r_s = -0.60$, $n = 12$, $p = .040$, two-tailed). Additionally, no significant correlations arose on the basis of depression, hostility, or positive affect.

**Performance on the PONS test**

Performance on the PONS test was measured as the proportion of correct responses, separately in terms of the video clips showing only the face, only the body, and overall, across both types of video clip (see Table 4.1). The results of independent-samples $T$-tests indicated that the scores of CCTV operators did not differ significantly from the scores of lay people on any of the three summary scores of the PONS test.

<table>
<thead>
<tr>
<th>PONS score</th>
<th>CCTV operators</th>
<th>Lay people</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>overall</td>
<td>0.71</td>
<td>0.06</td>
</tr>
<tr>
<td>body</td>
<td>0.72</td>
<td>0.11</td>
</tr>
<tr>
<td>face</td>
<td>0.73</td>
<td>0.08</td>
</tr>
</tbody>
</table>

*Table 4.1. Summary scores for CCTV operators and lay people on PONS test (i.e., proportion correct answers); (adapted from Blechko et al. (2008))*

**Correlations between sensitivity to the concealed firearm and performance on the PONS test**

Spearman’s correlations were conducted between sensitivity to the concealed firearm and each of the PONS test scores (face only, body only, overall) across participants. The analysis showed that the sensitivity to the concealed firearm did not correlate with the ability to read body language as assessed using any
of the summary scores of the PONS test (body only, $r_s = -0.03, n = 16, p = .917$; face only, $r_s = -0.26, n = 16, p = .340$; Overall, $r_s = -0.16, n = 16, p = .549$).

4.3. Discussion

It was assumed that the CCTV operators, due to their training and work experience, would be able to identify a concealed gun bearer more accurately than people without any training or experience, the current study examined the potential influence of expertise on the performance on the task of detection concealed firearm carrying. The main finding of the present study was that performance on the task of recognition of concealed gun carrying did not vary significantly between CCTV operators and lay people. Thus, it appears that training, work experience, and expertise did not improve performance on this particular task.

According to Diffley and Wallace (1998) training in general cannot provide a CCTV operator with full competence. In terms of the detection of someone carrying a concealed firearm it is a rare real life occurrence for a CCTV operator to come across and so their training will not encompass this fully. The results of the current study seem to support this statement. It showed that the task of detecting concealed firearm carrying is equally difficult for both trained CCTV operators and lay people. This finding is to some degree in accordance with the results of prior research into the abilities of CCTV operators and lay people to detect mal-intent via CCTV (Troscianko et al., 2004) that demonstrated that both groups are able to achieve the same level of detection in this task. Nevertheless, the study of Troscianko et al. (2004) did not consider the situations when the antisocial human behaviour included carrying firearms. The current results may therefore suggest two potential inferences. Firstly, that the existing CCTV operator training techniques are not completely applicable for this specific kind of task. It may indicate that in those training techniques there is a lack of specific methods needed to enhance the rate of
detection of concealed firearm carrying. For instance, such methods could address more the awareness of potential cues related to concealed firearm carrying. Secondly, the results of the current study may indicate that, as has been suggested previously (e.g., Donald, 1999), the attention for the natural predisposition of a worker to perform a certain task (e.g. spotting a bearer of concealed firearm) is necessary to enhance operators' performance.

The current study’s results might suggest that the task of concealed firearm carrying detection relies upon fundamental human abilities in visual perception and cognition as there was little difference between the trained and the naïve participants. Furthermore, the results of the current study indicates that there was a tendency for both CCTV operators and lay people to mistake a concealed innocuous object for a concealed firearm as their performance was below chance. One of the possible reasons for this finding is that the observers’ task in the current study placed great demand on discrimination of details throughout watching a large amount of repetitive video footage, which required both continuous concentration and detailed visual attention. In other words, the task was highly demanding in terms of perceptual and cognitive load and could result in fatigue, which in turn may have affected and possibly deteriorated the performance of participants. Fatigue typically occurs after some kind of prolonged visual activity, and is manifest by a decline in visual performance or an increase in visual discomfort, or both (Megaw, 1995). According to Montgomery, Montgomery and Guisado (1995) cognitive fatigue can be defined as the unwillingness to continue performance of mental work in alert and motivated subjects. With regard to the current research, presenting participants with a smaller number of video clips in the next studies should reduce any effect of fatigue on the overall performance in these tasks. In the current experiment a large number of trails were required in order to obtain the necessary level of statistical power and confidence level of detecting of cues associated with a firearm carrying.

In addition, even though there was a tendency for CCTV operators to demonstrate higher levels of sensitivity to firearms, this effect was medium-sized and approached statistical significance. This could mean that the
experiment may have been underpowered. Perhaps the use of a larger sample may have revealed a significant difference.

Other results of this study reveal that observers' decisions appeared to have been associated with the level of anxiety experienced by the surveillance target. The number of times a surveillance target was deemed to be carrying a concealed firearm was related to their level of anxiety when not carrying a firearm, as well as to the size of the influence of carrying a firearm on their level of anxiety. Additionally, a correlation between the number of times a surveillance target was deemed to be carrying a concealed firearm and their level of anxiety when they were carrying a firearm tended towards significance. Taken together, these results suggest that the mechanism by which the decision regarding whether or not a surveillance target was carrying a concealed firearm may have involved an interpretation of the surveillance target’s affective state.

These results imply that anxiety experienced by surveillance targets seemed to have the greatest influence on the number of times that an individual was deemed to be carrying a concealed firearm, across the five emotional states considered amongst the surveillance targets (i.e., anxiety, depression, hostility, positive affect and sensation seeking). It may be concluded that anxiety was the most influential affective state with respect to observers’ judgments relating to the carrying of a firearm. This leads to the assumption that anxiety might be the most visually apparent of the five types of affect considered. Given that there was no significant relationship between an observer’s firearm detection sensitivity and the size of the influence of carrying a firearm on a surveillance targets’ levels of anxiety, it can be assumed that an interpretation of the level of anxiety in a surveillance target does not support reliability in the detection of concealed firearms. In addition, it was found that the size of the influence of the firearm on the levels of sensation seeking experienced by surveillance targets was correlated negatively with firearm detection sensitivities, across surveillance targets. This finding suggests that visual cues related to sensation seeking could be used successfully as an indicator of the presence of a concealed firearm.
These results imply that those engaged in a surveillance task might use visual indicators of the affective state of a surveillance target to make a decision as to whether or not that individual is carrying a concealed firearm. The levels of concealed firearm detection sensitivity were consistently below zero, for both CCTV operators and lay people. Therefore, irrespective of any correlations between the affective state of a surveillance target and firearm detection sensitivity, it cannot be inferred that the use of visual indicators of affective state only would support reliability in the detection of concealed firearm carrying.

With respect to the measurement of the non-verbal behaviour decoding abilities of participants, the results show that there was no relationship between sensitivity to non-verbal cues and sensitivity to a concealed firearm amongst CCTV operators or lay people. There are a number of potential explanations for this finding. The PONS test might not capture the same aspects of body language reading as those used in the detection of a concealed firearm, the PONS test may therefore not be sensitive enough (Meiran, Netzer, Netzer, Itzhak and Rechnitz, 1994). The detection of a concealed firearm proved too difficult, and because of this, the detection of a concealed firearm in the present task does not rely significantly upon ability to read body language. The difficulty of the task could reduce the potential for detecting a relationship between the ability to read body language and the ability to detect a concealed firearm (a floor effect). The overall low levels of performance on the task of detection concealed firearm carrying would suggest that the latter explanation might be the main factor.

4.4. Conclusion

Prior research has shown that it is possible to use CCTV to detect lawlessness in a surveillance target on the basis of the immediate precursors of overtly violent behaviour (Troscianko et al., 2004). In the present study it was hypothesised that, when the surveillance task is to detect a firearm concealed
under the clothing of a surveillance target, there may be a reliance on more subtle visual cues such as non-verbal cues related to affective state.

The results of the present study showed that higher levels of anxiety in a surveillance target seem to be the most prominent cue that was used by observers in the decision process about the presence of a concealed firearm. However, anxiety levels did not appear to support reliable concealed firearms detection. Rather, sensation seeking in the surveillance target allowed a better differentiation between those who were carrying concealed firearms and those who were carrying an innocuous object. Even so, performance on the task of detection of a concealed firearm was consistently below chance. One of the possible reasons for this might be the fact that specific perceptual and cognitive fatigue could occur during the task that may have involved a high level of visual and cognitive effort due to a large number of video clips the participants were requested to watch, which in turn impaired their performance on this task. In order to reduce the effects of visual and cognitive fatigue in the following studies fewer video clips were used.

In the present study, it has been found that the decision regarding whether or not a surveillance target is carrying a concealed firearm is related to the affective state of that surveillance target. It is inferred that this phenomenon relates to the use, possibly at a subconscious level, of visual indicators of another’s emotional state. In the further experiments it remains to be determined whether visual cues related to affective states may influence the decision making process. For instance, the visual cues may operate overtly at a conscious level, or they may operate covertly at a subconscious level (Darker, Gale, Ward, Blechko and Purdy, 2007). Although the main responsibility of surveillance operators is to detect human mal-intent, their responsibilities are not limited to this task (Gill et al., 2005). Moreover, the task does not necessarily include constantly looking for people who are carrying an illegal firearm. Perhaps the detection of concealed firearms might be addressed using another, more indirect methodology. In the subsequent experiment, rather than asking participants to detect whether or not a surveillance target is carrying a firearm, they could be required to watch the video footage without being
informed explicitly about the presence of firearms concealed on the bodies of surveillance targets. Instead, the task of the observers is then to identify the affective states of surveillance targets. Since the participants are unaware of the ‘true’ task they are performing (i.e. gun carrier detection task), it is possible that the detection of a firearm carrier by an observer may be performed at a subconscious level. Furthermore, by applying this methodology additional information about the relationship between the detection of surveillance targets with a concealed firearm and the identification of their affective state can be gained. The next chapter will describe and discuss more extensively the study that includes this change in methodology.
CHAPTER 5: EXPERIMENT THREE

5.1. Introduction

The foregoing study described in Chapter 4 examined the potential influence of expertise by comparing the performance of real CCTV operators with that of lay people on the task of detection of a concealed firearm carrier from CCTV footage. It was assumed that the CCTV operators, due to their training and work experience, would be able to identify a concealed gun carrier more accurately than people without any training or experience. However, the results showed that task of differentiation between an individual with a concealed firearm and an individual with a concealed innocuous object was did not show significant differences between CCTV operators and lay people. Both groups were unable to discriminate between a concealed firearm bearer and bearer of concealed innocuous object. Thus, the performance was found to be less influenced by training and less related to the decoding abilities of participants than expected. This finding was in accordance with a previous proposition made by other researchers (e.g., Troscianko et al., 2004; Darker, Gale, Blechko, and Whittle, 2009) that the task of detecting precursors of human mal-intent relies upon fundamental human abilities in visual perception and cognition. In addition, although there was no effect of expertise found in Experiment Two, its results showed that the decision regarding whether or not a surveillance target is carrying a concealed firearm might be related to the estimation of the affective state of the target.

In Experiment Two (see Chapter 4) it had been assumed that in relation to surveillance work, the perception of surveillance targets’ affective state might underlie the decision whether or not this surveillance target is carrying a concealed firearm. The results showed that further investigation on this issue was needed. Since it was found that both groups, professionals and lay people,
performed on the task with the same level of performance, the subsequent experiments reported in this thesis only used lay people as participants.

One of the aims of the research reported in this thesis was to focus upon determining whether visual affect recognition from non-verbal behaviour can support the detection of concealed firearms, via CCTV. If it can then this is important as it posits an approach whereby computer software within the overarching MUDUSA project can be developed to identify such non-verbal behaviour in an automated fashion so as to aid the CCTV operator. Affect detection, by itself, is not necessarily a realistic real world task for general CCTV operations. Therefore, another attempt was undertaken to investigate whether this ability is more highly developed in people more sensitive to non-verbal cues. While in the previous study the relationship between the scores on the PONS test and the performance on the task of detection a firearm carrier was examined, this time the decoding abilities of observers was investigated in direct relationship with decoding of affective state of surveillance targets. This allows additional information to be gained about whether the PONS test has the potential to be a suitable measure of sensitivity to non-verbal behaviour in relation to the detection of concealed firearms carrying.

The following section describes a study which was based mainly on the literature review discussed earlier and on the findings from Experiment Two. The aim of the study was to investigate whether observers were able to perceive differences in the affective states of people who were and who were not carrying concealed firearms, as judged by monitoring staged CCTV footage. Additionally, similarly to Experiment Two, the test of affect decoding abilities was performed in order to examine the relationship between individual differences in sensitivity to non-verbal cues and performance on the affect detection task.
5.2. Experiment Three

5.2.1. Introduction

As demonstrated previously (Chapter 4, section 4.2), a person’s emotional state can be conveyed through their non-verbal behaviour, which in turn can be picked up and interpreted by an observer. In accordance with the evidence from the literature (e.g., Hales et al., 2006; Anderson et al., 1998), the results obtained by Experiment One have shown that handling a real firearm produces a ‘weapons effect’ involving a negative affective response in the carrier of the firearm. To be able to investigate whether the conscious differentiation between the affective states of people is related to detection of the surveillance targets with a concealed firearm another methodology, as compared to Experiment Two, has been used. Considering the ability of a weapon to increase the accessibility of aggressive thoughts, in particular, by ordinary people (Anderson et al., 1998) it was decided to exclude the effect of knowledge about the presence of a firearm in the footage. This was done by asking participants to identify an affective state through watching the footage of people walking with concealed firearms and innocuous matched concealed objects, used in Experiment Two, without informing them about the presence of the firearms in the footage. By applying this method to the task of detecting surveillance targets with a concealed firearm, the underlying effect of the firearm on the target’s affective state can be studied. The task used in this experiment (i.e. the task of detection of only the affective state of surveillance targets) would allow imitating certain real life situations. Examples of context in which perception of affective state of surveillance targets might be important are VIP protection, surveillance of some public places such as banks or hotels in order to recognize atypical events, or surveillance targeting to detect passengers in airplanes who might form a threat to others or themselves.

Furthermore, the PONS test of affect decoding abilities was again used to determine whether a relationship could be established between general sensitivity to non-verbal cues and performance on the task of recognising the
affective state of surveillance targets. It is known that rating someone else’s affective state can be vulnerable to the rater’s own traits. The possibility of the appearance of ‘projection error’ in the peer-ratings (i.e., the tendency to rate others high or low on the traits on which they rate themselves as high or low) can be an influential factor in biasing the observer’s ratings (Epkins, 1994). The traits of observers in the present study were therefore measured prior to the experiment using the MACCL-R Trait questionnaire.

To summarize, the present study investigates the abilities of observers to read the affective states of people viewed via CCTV; some of these people were bearing a concealed firearm, and from Experiment One they were known to have felt more aggression whilst carrying the firearm. The following research questions were formulated:

*Is it possible to differentiate between surveillance targets when they are carrying a concealed firearm and when they are carrying a concealed innocuous object, in the mock CCTV footage, based on an estimation of the surveillance target’s affective state? (Affect Detection task: AD task).*

*How does the estimation of a gun carriers’ affective state by observers match the ratings of the affective state as made by the surveillance targets themselves?*

*Whether or not the decoding ability (i.e., sensitivity to bodily and facial expressions) of observers was related to their performance on AD task, obtained by calculating the absolute difference between perceived by observers and experienced by surveillance targets’ affective states?*
5.2.2. Method

5.2.2.1. Participants

Thirty-one undergraduates and postgraduates students from Loughborough University (20 male and 11 female) volunteered to take part. The age of participants varied between 20 to 35 years, (mean age 25.2, $SD = 4.7$). None had previous experience with surveillance work or in spotting criminal behaviour in general. All participants had normal or corrected-to-normal vision. They were naïve to the purpose of the study and were paid for their participation.

5.2.2.2. Materials

The stimuli consisted of PONS test video clips (Rosenthal et al., 1979), mock CCTV video clips made in the Experiment One (see Chapter 2, section 2.3. for a complete description of this aspect of the experiment) and the Multiple Affect Adjective Check List - Revised (MAACL-R; Lubin and Zuckerman, 1999) questionnaire.

The mock CCTV footage, the PONS test and the MAACL-R questionnaire, were all presented in the centre of a 14-inch laptop computer (with operating system Microsoft Windows XP, version 2002) colour monitor, with frame size of 16 cm in height and 20.8 cm in width.

The length of each PONS video fragment was 3.5 seconds. In total 144 mock CCTV video clips with 12 different surveillance targets in two conditions (concealed firearm present; concealed innocuous object present) were shown, each time in a group of six video clips. The full length of each batch of six video clips was 12 seconds; each short CCTV video fragment was 2 second long. Additionally, the paper-and-pencil version of the MAACL-R Trait was used.
Affective State measures

The measurement of participants’ estimation of surveillance targets’ affective state in the staged video footage was performed using the MAACL-R State questionnaire (described in Chapter 2, section 2.4.2.).

The instruction for the MAACL-R State form was as follows: “Tick the adjectives that describe the mood of the person you have just seen in the immediately prior video clips”. In order to measure the traits of the participants, the MAACL-R Trait was used. When subjects were given the MAACL-R Trait form (prior to the experiment) they were instructed to check those mood adjectives that described how they “generally feel”, or how they have “generally felt over the past month”.

Measure of sensitivity to non-verbal cues

In order to measure the participants’ sensitivity to non-verbal cues the Face and Body profile of nonverbal sensitivity was used, which is a shortened version of the full PONS test (Rosenthal et al., 1979). The PONS face score and the PONS body score were used separately in the analysis in order to obtain information on decoding skills. A full description of the PONS test is given in Chapter 2, section 2.6.

Design

A full-factorial design with repeated measures was used to vary the Type of Object Concealed (i.e., firearm versus innocuous object). Every experimental trial consisted of a batch of six video clips (see Table 5.1), either with a matched object, a concealed bottle (i.e., Innocuous Object trial) or with a concealed firearm (i.e., Gun trial).
Table 5.1. Overall number of clips used in the study in every kind of trial (Innocuous Object and Gun) with different kind of carrying object (i.e., three types of bottles and three types of firearms), for each surveillance target.

Only the concealed staged video clips were used in this experiment (see Chapter 2, Section 2.3). There was in total 144 video clips with 12 different surveillance targets in two conditions (i.e., Type of Object Concealed; firearm versus innocuous object). Each of the 12 surveillance targets was showed separately, in one experimental trial; the order of Gun and Innocuous Object trials was counter-balanced across participants.

5.2.2.3. Procedure

After the participants arrived, they were seated approximately 60 cm in front of the computer screen. Prior to the experiment all participants were given a form with a MAACL-R Trait questionnaire and they were asked to describe how they generally felt. After they completed the questionnaire the experiment began. This consisted of two sessions. Each session started with a practice so as to familiarise participants with the task. In the first session the participants had to watch short video clips from the PONS test and after each clip they were
instructed to pick one of the two statements that best described the situation in which the emotion was portrayed. The participants were asked to choose the answer as quickly and as accurately as possible. The first session took approximately 5 minutes.

In the second session the observers were asked to watch a batch of six short video clips of a surveillance target and then use the response form (i.e., MAACL-R State questionnaire) shown on the screen afterwards to record how they thought the person in those video clips was feeling. To do this they needed to tick all the adjectives that described the mood of this person. The time needed to fill in the questionnaire (both, the paper-and-pencil version and the version presented on the computer screen) was not limited, it took approximately 2-3 minutes to complete. The second session was 45-50 minutes in length.

5.2.3. Results

Differentiation between carriers of a concealed firearm and carriers of a concealed innocuous object

In order to investigate whether or not observers were able to discriminate between the affective states of surveillance targets in the two conditions, ‘Innocuous Object’ and ‘Gun’ a repeated measures Analysis of Variance (ANOVA) was used, with the Type of Object Concealed (Innocuous Object versus Gun) as the within-subjects factor. There was a significant effect of the Type of Object Concealed on the perceived level of anxiety ($F(1,30) = 8.192, p = .008$) and dysphoria ($F(1,30) = 6.885, p = .014$) for the MAACL-R questionnaire. The mean score on these scales in the Innocuous Object condition was significantly higher than the mean score in the Gun condition (for all means and standard deviations see Table 5.2).
Innocuous Object (Bottle)                   Gun (Firearm)  

<table>
<thead>
<tr>
<th>Scales</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>MSE</th>
<th>p</th>
<th>η\textsuperscript{2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
<td>0.76</td>
<td>0.61</td>
<td>0.58</td>
<td>0.46</td>
<td>8.192</td>
<td>0.62</td>
<td>.008*</td>
<td>0.214</td>
</tr>
<tr>
<td>Depression</td>
<td>0.85</td>
<td>0.86</td>
<td>0.70</td>
<td>0.62</td>
<td>3.790</td>
<td>0.93</td>
<td>.061</td>
<td>0.112</td>
</tr>
<tr>
<td>Hostility</td>
<td>0.67</td>
<td>0.87</td>
<td>0.69</td>
<td>0.82</td>
<td>0.155</td>
<td>0.88</td>
<td>.697</td>
<td>0.005</td>
</tr>
<tr>
<td>Dysphoria</td>
<td>2.28</td>
<td>2.09</td>
<td>1.95</td>
<td>1.71</td>
<td>6.885</td>
<td>0.21</td>
<td>.014*</td>
<td>0.187</td>
</tr>
<tr>
<td>Sensation</td>
<td>4.10</td>
<td>0.68</td>
<td>4.18</td>
<td>0.61</td>
<td>0.863</td>
<td>0.12</td>
<td>.360</td>
<td>0.028</td>
</tr>
<tr>
<td>Sensation</td>
<td>4.10</td>
<td>0.68</td>
<td>4.18</td>
<td>0.61</td>
<td>0.863</td>
<td>0.12</td>
<td>.360</td>
<td>0.028</td>
</tr>
<tr>
<td>Hostility</td>
<td>4.10</td>
<td>0.68</td>
<td>4.18</td>
<td>0.61</td>
<td>0.863</td>
<td>0.12</td>
<td>.360</td>
<td>0.028</td>
</tr>
<tr>
<td>Positive Affect</td>
<td>0.85</td>
<td>0.80</td>
<td>1.27</td>
<td>1.51</td>
<td>6.453</td>
<td>0.42</td>
<td>.019*</td>
<td>0.177</td>
</tr>
<tr>
<td>PASS</td>
<td>4.95</td>
<td>1.28</td>
<td>5.45</td>
<td>1.94</td>
<td>7.231</td>
<td>0.54</td>
<td>.012*</td>
<td>0.194</td>
</tr>
</tbody>
</table>

Table 5.2. Comparison of mean scores of observers on all scales and subscales of MAACL-R questionnaire in the Innocuous Object and Gun conditions (repeated measures ANOVAs with df = 1/30). * p < .05

This contrasted with the surveillance targets’ own ratings which revealed that their levels of anxiety and dysphoria were higher in the Gun condition than in the Innocuous Object condition. Only the differences in the mean scores on the Hostility scale were in the same direction as the scores of the surveillance targets (see Figure 5.1), however this effect did not reach statistical significance.

There was a significant effect of the Type of Concealed Object on the observers’ scores on Positive Affect and PASS scales, \( F (1,30) = 6.453, p = .019; F (1,30) = 7.231, p = .012, \) respectively). The differences in mean scores show that the observers rated the mood of the surveillance targets in the Gun condition as more positive than the mood of surveillance targets in the Innocuous Object condition. The surveillance targets own ratings of their mood did not differ with the Type of Concealed Object.
Figure 5.1. Mean scores on the different scales of MAACL-R questionnaire (Anxiety, Depression, Hostility, Dysphoria, Positive Affect, Sensation Seeking and PASS) in Innocuous Object (i.e., surveillance target with a bottle) and Gun (i.e., surveillance target with a firearm) conditions, which were given by observers in their rating of the surveillance targets’ affective state.

Performance on Affect Detection task

In addition, the absolute differences between the observers’ scores and the scores of surveillance targets (i.e., modulus, both presented as a percentage of the maximum possible score) were calculated in order to attain more insight into how well the observers were performing on the AD task.

The analysis was carried out separately for the negative affective scales of the questionnaire (i.e., Anxiety, Depression, Hostility and Dysphoria). Analysis of Variance with the Type of Object Concealed (Innocuous Object versus Gun) as a within subjects factor showed a significant main effect of the Type of Object Concealed for all negative scales of the MAACL-R questionnaire \( (F(1,30) = 218.350, p = .000, \) for Anxiety; \( F(1,30) = 22.721, p = .000, \) for Depression; \( F(1,30) = 46.852, p = .000, \) for Hostility; \( F(1,30) = 109.336, p = .000, \) for Dysphoria). This showed that in the Innocuous Object condition the differences between observers’ and surveillance targets’ scores
were significantly lower than these differences in the Gun condition (see Table 5.3). This means that the observers were better (more accurate) at judging each surveillance target’s negatively valenced affective states when they were carrying an innocuous object than when they were carrying a firearm.

When analysing only the differences in scores on the positive affective scales of the MAACL-R questionnaire (i.e., Positive Affect, Sensation Seeking and PASS), the results showed that the differences between observers’ and surveillance targets’ scores were significantly higher in the Innocuous Object condition ($F(1,30) = 48.103$, $p = .000$, for Positive Affect; $F(1,30) = 32.772$, $p = .000$, for Sensation Seeking; $F(1,30) = 20.142$, $p = .000$, for PASS). This means that the performance on the AD task was more accurate when the surveillance targets were carrying a firearm, with the exception of the scores on the Sensation Seeking scale (see Table 5.3). The scores on this scale in the Gun condition were significantly lower than in the Innocuous Object condition which indicates that the observers were more accurate in the detection of sensation seeking in a surveillance target when they were carrying an innocuous object than when they were carrying a firearm.

<table>
<thead>
<tr>
<th>Scales</th>
<th>Innocuous Object (Bottle)</th>
<th>Gun (Firearm)</th>
<th>F</th>
<th>MSE</th>
<th>p</th>
<th>$\eta_p^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
<td>9.64</td>
<td>21.72</td>
<td>218.350</td>
<td>10.36</td>
<td>.000*</td>
<td>0.879</td>
</tr>
<tr>
<td>Depression</td>
<td>7.38</td>
<td>11.34</td>
<td>22.721</td>
<td>10.72</td>
<td>.000*</td>
<td>0.431</td>
</tr>
<tr>
<td>Hostility</td>
<td>6.66</td>
<td>9.76</td>
<td>46.852</td>
<td>0.043</td>
<td>.000*</td>
<td>0.61</td>
</tr>
<tr>
<td>Dysphoria</td>
<td>6.81</td>
<td>12.23</td>
<td>109.336</td>
<td>4.159</td>
<td>.000*</td>
<td>0.785</td>
</tr>
<tr>
<td>Sensation Seeking</td>
<td>15.98</td>
<td>19.50</td>
<td>32.772</td>
<td>5.85</td>
<td>.000*</td>
<td>0.522</td>
</tr>
<tr>
<td>Positive Affect</td>
<td>21.81</td>
<td>14.89</td>
<td>48.103</td>
<td>15.42</td>
<td>.000*</td>
<td>0.616</td>
</tr>
<tr>
<td>PASS</td>
<td>17.69</td>
<td>15.50</td>
<td>20.142</td>
<td>3.68</td>
<td>.000*</td>
<td>0.402</td>
</tr>
</tbody>
</table>

*Table 5.3. Comparison of mean differences in scores (%) of observers and surveillance targets on all scales and subscales of MAACL-R questionnaire in the Innocuous Object and Gun conditions (repeated measures ANOVA with df = 1/30).* $p < .05$
To summarize, the observers were able to discriminate between the affective states of surveillance targets when they were carrying a firearm and when they were carrying an innocuous object. Furthermore, the results imply that whilst carrying a firearm the surveillance targets tend to appear to observers in such a way that their affective state was interpreted as one with a positive valence. Moreover, when surveillance targets were carrying a firearm, only their level of positive affective states (with the exception of the level of sensation seeking) was detected by observers more accurately than other affective states.

*Relationship between the measures of decoding ability (i.e., sensitivity to bodily and facial expressions) and the performance on AD task*

Partial correlation analysis was performed in order to investigate whether the observers’ sensitivity to non-verbal cues was related to their performance on AD task, controlling for their trait scores. Performing a partial correlation analysis allows the removal of possible influence of observers’ traits on the correlation between the decoding abilities and performance on AD task.

There was a significant positive correlation between the scores of observers on the PONS Face test and the differences in scores on the Anxiety scale between observers and surveillance targets in the Innocuous Object condition ($r = .487$, $n = 31$, $p = 0.047$). No other correlations between the scores of these two tests were found. It can be concluded that there was little relationship between the performance on the PONS test and the observers’ peer-rating differences between observers’ and targets’ rating of affective state and their performance on AD task.

At the beginning of this study, it was assumed that the traits of the observers may affect the ratings of the targets’ mood and may hinder these ratings in such a way that the scales of the questionnaire which describe emotional state similar to the emotional state of the observers receive higher scores than the scales which describe a different emotional state. Consequently, the overall observers’ rating of the mood of targets could be shifted in the direction of the observers’ own traits. However, after performing a
control on the effects of traits the results showed that there was no evident relationship between the traits of observers and their performance on mood detection task.

5.3. Discussion

Differentiating a surveillance target with a concealed firearm from other people is a highly complex task. Obvious visual cues (e.g., clothing characteristics) might be used by an observer in doing this, along with more subtle cues, such as non-verbal cues related to a certain affective state of the gun carrier (e.g., gestures, facial expressions or gait). Before investigating more closely which cues exactly the observers could use in such a task, the present study investigated whether observers, inexperienced in surveillance tasks, are able to perceive differences between the affective states of surveillance targets who were carrying either an innocuous object or a firearm. Additionally, the study examined the relationship between the observers’ sensitivity to non-verbal cues and their ability to differentiate between the affective states of surveillance targets.

Due to the experimental design (i.e., all the video clips showed the carriers of only concealed objects, either firearm or an innocuous object), and to the instructions received by the participants (i.e., they were not informed about the presence of firearm in the video clips) the observers were unaware that they were watching two groups of surveillance targets in the video clips (i.e., who were carrying a concealed firearm versus an innocuous object). The results showed that even without knowing about this experimental manipulation, the observers were able to discriminate, perhaps on a subconscious level, between the affective states of two groups of surveillance targets by attributing different descriptions of affective state to the targets. The differences in scores between the two conditions were significant on both negative and positive scales of the emotional state questionnaire.
Furthermore, when the targets were carrying a firearm, their affective state of positive valence, was more accurately detected by observers than their mood, associated with other (i.e., negative) scales of the MAACL-R. This indicates that while carrying a gun the surveillance targets displayed non-verbal behaviour which was interpreted by observers as an expression of an affective state with a positive affective valence. Moreover, the expression of positive affective states was detected more accurately than an affective state with a negative valence.

The results regarding the relationship between the measurement of sensitivity to non-verbal cues and the performance on the affect detection task showed that there was only modest evidence found for the assumption that the decoding ability of observers, as measured by the PONS test, related to their ability to recognize the affective state of people in the staged CCTV footage. The better decoding of facial expressions was surprisingly related to less correct detection of the level of anxiety in surveillance targets. Taken together, the results of both experiments (i.e. studies Two and Three) related to the use of the PONS test demonstrated that the investigated relationship between the measurement of sensitivity to non-verbal cues and the performance on both affect detection and gun carrier detection tasks showed too little evidence to be able to conclude that the PONS test can would predict the ability to either identify a firearm bearer or to recognize their affective state with the present mock CCTV footage. Therefore, the relationship between performance on PONS test and the performance on affect and gun carrier detection tasks was excluded from further examination in the subsequent studies.

The lack of the expected relationship between measurements of the observers’ decoding abilities and their performance on both tasks in the current research may be due to range of factors. For instance gender differences in non-verbal encoding abilities which may have influenced the way an affective state was expressed by only a female encoder in PONS test and only male actors in the staged CCTV clips (DePaulo, 1992). As mentioned in the previous chapter, the PONS test may therefore not capture the same aspects of reading body language as those used in the detection of concealed firearm carriers or
their affective state, and also it may not be sensitive enough (Meiran et al., 1994) for this kind of task. This subject and the implications for the future research of the findings related to the use of PONS test will be discussed in more details in the conclusion of the thesis.

In addition, the results of the present study showed the discrepancy between the surveillance targets’ self-ratings of their affective state and the peer-ratings of this state given by observers. As indicated by the results of Experiment One, the surveillance targets rated their own affective state as more negative whilst carrying a firearm than whilst carrying an innocuous object. The current study showed that although there was a noticeable agreement amongst observers regarding their estimation of the targets’ affective states, this rating varied in an opposite direction to the self-rated affective state reported by the surveillance targets themselves. This implies that despite the ability to differentiate between a firearm bearer and a bearer of an innocuous object mentioned earlier, the observers failed to recognize the ‘true’ affective state of bearers. In particular the reverse direction of those two ratings was an unexpected and rather remarkable finding.

What could be the cause of such discrepancies in assessed affective state with respect to carrying a concealed firearm? There might be several explanations for this. According to previous research on extraction of emotions from overall body movements and from gait in particular, emotions such as joy and anger, which are opposite in valence but both associated with high levels of arousal, could both produce the same visual cues (e.g., greater stride lengths and higher velocity; Cluss, Crane, Gross, and Fredrickson, 2006) and were linked to the same observers’ judgments, such as high movement activity, expansive movement and high movement dynamics (Wallbott, 1998). This could affect the observers’ judgment of the surveillance targets’ affective state in the current study, and may give an explanation why detected affective states were not congruent with the affective states reported by the surveillance targets.

Another possible explanation which could be considered here is the influence of the viewing angle on perception of human face and body and how
this affects the ability to discriminate between affective states. In the present study the images from only one viewing angle were used, whilst from previous research it is known that different angles of viewing faces and body movements result in a different degree of recognition of certain characteristics of the observed subject (Troje, Westhoff, Lavrov, 2005; Troje and Bulthoff, 1996). Furthermore, the experiments which were performed to study the identification of human actions and body postures viewed from different viewpoints (e.g., Daems and Verfaillie, 1999) suggest that the stored representations of human postures that mediate the identification of human actions and postures are viewpoint specific. Moreover, evidence showed that there is a general tendency for frontal views to lead to more consensual attributions of affective states to presented postures (Coulson, 2004).

According to some evidence from the literature, even from poor-quality CCTV images identification of target people was remarkably affected by obscuring peoples' heads, compared with a situation when their body or gait were obscured (Burton, Wilson, Cowan, and Bruce, 1999). In the present experiment, no attempt was made to explore which distinctive features (e.g., face versus body of the firearm bearer) were typical of the observed affective states of surveillance targets. This results in a lack of understanding about why the appearance of certain surveillance targets was perceived as more negative than another target's appearance, and which cues in particular were associated with this. In the studies presented next, more attention will be therefore paid to analysing which cues observers were using in their description of surveillance targets' affective state.

5.4. Conclusion

The present study showed that untrained observers are able to differentiate between the images of people walking with or without a concealed firearm, based on the affective state that they attribute to them and without knowing about the presence of different concealed objects. They were more accurate in
their detection of surveillance targets' positive affective state when the targets were carrying a concealed firearm object. However, a notable result was that the observers could not infer correctly the self-estimated affective states of the surveillance targets. Assuming that this could be due to the fact that the recognition of affective state is viewpoint dependent and that the frontal view might produce more correctly identification of affective state in people, the further investigation presented in the following chapter will be done by varying the viewing angle, i.e. by presenting images of surveillance targets from a frontal view (i.e., street-level perspective) to observers.

Furthermore, questions arise regarding the cues being used by observers and the cues displayed by surveillance targets; for instance, which produced the misinterpretations? To investigate this, along with an attempt to determine which visual cues observers use during the Affect Detection task, further studies were performed as discussed in the following chapters. The visual cues used in attempting to decode the non-verbal cues displayed by carriers of concealed guns will be inferred by means of a questionnaire and determined later empirically by eye-tracking.
CHAPTER 6: EXPERIMENT FOUR

6.1. Introduction

As outlined in the Introduction of this thesis, evidence exists documenting that people are adept at judging the emotional states of others on the basis of non-verbal cues such as facial expressions and body movements. In the absence of audio information, such cues might be seen as key to CCTV surveillance operatives in their efforts to deduce the intentions of surveillance targets. The prior work (see study One, chapter 3) which investigated the influence of gun carrying on the affective state of the carrier, showed that the gun carrying was associated with increased dysphoria in the gun bearer. Study Three presented in chapter Five aimed to assess the accuracy with which an observer can judge the affective state of another when viewed in CCTV footage from a typical, elevated, CCTV-like perspective. The findings of this study suggest that untrained observers are able to differentiate between individuals walking with or without a concealed firearm, based on the affective state that they attribute to them. The surprisingly result was, however, that this attribution of affect was not just incongruent with the self-ratings of surveillance targets, but also in the opposite direction compared to the affect ratings made by the surveillance targets. The observers attributed a greater degree of negative affect to the surveillance targets in the condition when the targets were concealing an innocuous object than to the surveillance targets in the condition when they were concealing a firearm. Based on the evidence from existing research on viewpoint dependence (Daems and Verfaillie, 1999; Coulson, 2004) it was assumed that the possible explanation for this result could be the fact that the task of affect recognition in study Three was made from the particular view (i.e., CCTV–like perspective) that reveals potentially less easily accessible information about the affective state of individuals in the imagery than, for example, would a ‘street-level’ perspective (i.e., the camera placed at head height), and a frontal view.
Optimal CCTV camera placement has been discussed and emphasized by a number of authors (e.g., Bodor, Drenner, Schrater, and Papanikolopoulos, 2007; Fiore, Fehr, Bodor, Drenner, Somasundaram, and Papanikolopoulos, 2008). The variations in the placement of surveillance cameras (e.g. too high or too much to the left or right) are believed to affect the degree of effectiveness of surveillance systems (Bodor et al., 2007). An ad-hoc “intuitive” placement of multiple cameras viewing a scene is known not to observe human activities (e.g., pedestrians’ activities) in an optimal manner (Fiore et al., 2008). According to the CCTV Operational Requirements manual 2009 (Cohen et al., 2009) it may be advisable to place the surveillance cameras at head height where suspect identification is the main priority, as ceiling mounted cameras may not be able to provide a full view of the suspect’s face. Although the positioning of cameras above head height would be useful due to such factors as the need for physical protection, e.g. from weather or human interference, this may compromise the field of view and make facial identification more difficult. As mentioned previously (see Chapter Five) some experimental work that was performed to study the identification of human actions, body postures and affective state viewed from different viewpoints, implies that its identification is viewpoint specific (e.g., Daems and Verfaillie, 1999; Coulson, 2004). Thus, previous work suggests that the viewpoint and the camera placement are important issues to consider when performing research related to the visual detection of concealed firearm carrying through CCTV.

Accordingly, the experiment\(^2\) presented in the current chapter will investigate the topic of affect recognition of those who are concealing firearms or innocuous objects using video-footage of the same surveillance targets but when viewed from a street-level perspective (i.e., placed at head height) in order to make a comparison of the results with the results obtained in the previous experiment. Furthermore, the visual cues used by the observers whilst

\(^2\)The paper that described this study ‘The Role of Emotion Recognition from Non-Verbal Behaviour in Detection of Concealed Firearm Carrying’ has been published by Blechko, A., Darker, I.T., Gale, A.G. in the Proceedings of the Human Factors and Ergonomics Society, 53\(^{rd}\) Annual Meeting -2009.
performing the task will be explored by means of their self-reports obtained from a questionnaire.

Similar to study Three, the participants in the current study watched the footage and estimated the affective state of surveillance targets without knowledge about the presence of a firearm. Observers’ responses were later matched to the self-ratings of the surveillance targets. Furthermore, two Likert-scale questionnaires were developed and used to gain information about characteristics of the targets’ movement pattern (i.e., cues). It was believed that this would help to understand how exactly carrying a firearm affects the appearance of the bearer to the observer in terms of cues related to the bearer’s affective state. The questionnaires might provide data on which cues convey information needed to identify the affective state in a firearm and non-firearm bearer. As mentioned in Chapter 2, section 2.7.1., the questionnaire was based on previous work that assessed the identification of emotions and associating this with visual characteristics of human movements (e.g., Montepare et al., 1987; Montepare et al., 1999; Paterson et al., 2001; Pollick et al., 2001). One of the results showed for example that anger in human movements was rated as the most ‘jerky’ whereas movements presenting sad and neutral affective states were rated as the most ‘smooth’ (Montepare et al., 1999). Furthermore, the arm movements, for example, were a predictor of fear, whereas a lower degree of fear was associated with a higher degree of arm movements (Meijer, 1989). For the description of the questionnaire see Appendix 4. In the current study the following research questions were formulated:

Is it possible to differentiate between surveillance targets when they are carrying a concealed firearm and when they are carrying a concealed innocuous object, in mock CCTV footage when viewed from a street-level perspective, based on an estimation of the surveillance target’s mood? (Affect Detection task: AD task).

Which visual cues do the observers use in order to perform this task? Which visual cues are related to the condition when the firearm was present, in particular?
Is the level of performance on the AD task related to the use of certain visual cues?

6.2. Method

6.2.1. Participants

Thirty-one undergraduate and postgraduate students from Loughborough University participated in the experiment (14 female and 17 male). Their mean age was 24.6 years (range 18 - 34 years, $SD = 4.74$). None of the participants had previous experience with surveillance work or in spotting criminal behaviour in general.

6.2.2. Materials

Apparatus
Stimulus presentation and data collection were automated using a computer (Toshiba Tecra, Toshiba Ltd.; Microsoft Windows XP version 2002; Genuine Intel® Centrino DuoTM T2600 processor 2.16 GHz; 994 Mhz motherboard) and controlled by a program developed in-house.

Mock CCTV footage
In the methodology section (Chapter 2, section 2.3.) it was described how the idealised, ‘staged’ CCTV video footage of twelve male students, each carrying either a firearm or a matched innocuous object, was generated. For the purpose of the current study the footage made from a different viewpoint (i.e., frontal) was used, which was viewed by the observers from a street-level perspective. Each surveillance target was filmed separately; the order of gun and matched object filming was counter-balanced across participants, facilitating an analysis of the effect of carrying a gun compared to carrying an innocuous object. In the current study 22 video clips with 11 different surveillance targets in two conditions
(surveillance target carries a concealed firearm; surveillance target carries a concealed innocuous object) were employed. The video clips of one surveillance target were excluded from the study due to the irregularities of the video image. Each video clip appeared in the middle of the computer screen in the form of a rectangle, the size of which was 23 cm in height and 18 cm in width. The participants were shown a person walking towards the observer, the range of the size of the walking person varied between 9.3 cm (at the start point of walking path) and 12.7 cm (at the end point of walking path). The duration of each video clip was 2 seconds.

**Affect Detection task**

The levels of positive and negative affect were assessed by each observer for each surveillance target, both whilst carrying a firearm and whilst carrying the innocuous object, using the Multiple Affect Adjective Check List – Revised, State (MAACL-R; Lubin and Zuckerman, 1999). This was achieved within an experimental design that was counterbalanced for order of exposure to the firearm and innocuous object. The questionnaire consisted of two scales: Dysphoria (subscales: Anxiety, Hostility and Depression); PASS (subscales: Positive Affect and Sensation Seeking). For more detailed description of the questionnaire see Chapter 2, section 2.5.2.)

**Cue-detection questionnaires**

The cue-detection questionnaires aimed to reveal characteristics of the targets’ movement pattern as perceived by the observers. Movement pattern was specified by eight characteristics (*i.e.*, cues), namely: length of stride, degree of arms’ swinging, degree of heavy footedness *(related to surveillance targets’ gait)*; degree of fluidity, degree of rigidity, degree of speed, and degree of exaggeration *(related to the surveillance targets’ overall movements)*. Furthermore, the participants were asked to indicate how they gained an impression of the moods of the targets by means of rating four statements about whether or not they were looking at the gait, the posture, the facial expressions or the position of the arms of surveillance targets on five-point
scale \((i.e., \text{‘not at all’}, \text{‘rarely’}, \text{‘sometimes’}, \text{‘often’} \text{ and ‘all the time’})\). For the complete description of the questionnaire see Chapter 2, section 2.7.1., or the Appendix 4.

**Design**

A full-factorial design with repeated measures was used in this study to vary the type of staged video clips (Innocuous Object versus Gun). Every experimental trial consisted of one video (see Table 6.1), either with a matched object, a concealed bottle \((i.e.,\text{ Innocuous Object trial})\) or with a concealed firearm \((i.e., \text{ Gun trial})\).

<table>
<thead>
<tr>
<th>Concealed</th>
<th>Trial Innocuous Object (Bottle)</th>
<th>Trial Gun (Firearm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking towards the camera</td>
<td>1 x 1 litre (Total 11 clips)</td>
<td>1 x Glock (Total 11 clips)</td>
</tr>
</tbody>
</table>

Table 6.1. Overall number of clips used in the study in each kind of trial (Innocuous Object and Gun) with different kind of carrying object \((i.e., \text{bottle and firearm})\), for each surveillance target.

In this experiment only staged video clips involving concealed objects were used. There were, in total, 22 video clips with 11 different surveillance targets in two conditions (Innocuous Object and Gun). Each of the 11 surveillance targets was shown separately, in one experimental trial; Innocuous Object and Gun trials were shown in separate blocks of stimuli. The order of Innocuous Object and Gun blocks was counter-balanced across participants.
6.2.3. Procedure

The task of participants consisted of two parts. In the first part the observers were required to perform the AD task. They watched the surveillance target in twenty-two video clips (involving a full frontal view of the person walking towards camera), shown in a pseudo-randomised order. The Innocuous Object and Gun trials appeared in separate blocks. In order to record how participants thought the person in the prior video clip was feeling, the MAACL-R questionnaire was shown after each video clip. In the second part the observers were asked to answer the cue-detection questionnaire in order to describe the overall movement of surveillance targets using a number of characteristics of the targets’ movement pattern, and to indicate how observers gained an impression of the affective states of the surveillance targets in the clips.

6.3. Results

*Differentiation between the perceived affective states of concealed firearm carriers and carriers of a concealed innocuous object*

A repeated measures ANOVA, with the Type of Object Concealed (firearm versus innocuous object) as the factor was performed. An α-level of .05 was applied to tests of statistical significance.

A significant effect of Type of Object Concealed on the perceived level of dysphoria ($F(1,30) = 5.860, p = .022$) in the surveillance target was found. The differences in mean scores show that the observers attributed a greater degree of dysphoria to the affective states of the surveillance targets in the firearm condition than to the affective states of the surveillance targets in the innocuous object condition. This was congruent with the self-ratings of surveillance targets (see Figure 6.1).
Figure 6.1. The surveillance targets’ affective state, estimated by observers. The surveillance targets’ affective state is represented by the mean scores on the scales of MAACL-R in Innocuous Object (i.e., carrying a bottle) and Gun (i.e., carrying a firearm) conditions; (adapted from Blechko et al. (2009))

**Performance on Affect Detection task**

A repeated measures ANOVA, with the Type of Object Concealed (firearm versus innocuous object) as the factor showed a significant effect for Type of Object Concealed on the performance of observers in the AD task ($F(1,30) = 97.750, \ p = .000$, for Dysphoria; $F(1,30) = 38.993, \ p = .000$, for PASS). In the Innocuous Object condition, the differences between the observers’ and targets’ scores on negative affect scales were significantly lower than these differences in the Gun condition. This means that the performance of observers in the AD task was more accurate when the surveillance targets were carrying an innocuous object.
Relationship between performance on the Affect Detection task and the visual cues used by observers in this task

Pearson correlations were performed in order to investigate the possible relationship between the performance of observers on the AD task and the visual cues they used whilst performing the task. When the firearm was present a number significant correlations were found. There was a significant positive correlation between the perceived level of dysphoria and the degree to which the gait and the posture of the target were used by observers in order to gain an impression of the targets’ affective state ($r = 0.358$, $n = 31$, $p = 0.048$; and $r = 0.359$, $n = 31$, $p = 0.047$, respectively).

Furthermore, a significant negative correlation was found between the level of perceived dysphoria and the degree to which the facial expressions of the target were used by observers in order to gain an impression of the targets’ affective state ($r = -0.370$, $n = 31$, $p = 0.041$).

Additionally, there was a significant positive correlation between the performance on the task of estimation of dysphoria level (i.e., the difference between the observers’ and surveillance targets’ estimation of dysphoria) and the degree of perceived arms’ swinging in the movements of surveillance targets ($r = 0.355$, $n = 31$, $p = 0.050$).

6.4. Discussion

The results of this experiment showed that the observers were able to differentiate between surveillance targets when carrying a concealed firearm or an innocuous object on the basis of inferred affective state. Compared to the findings of experiment Three the level of dysphoria attributed to surveillance targets was estimated correctly. Regarding the visual cues the observers were using to perform this task, it was found that the gait and the posture of surveillance targets were related to the estimated level of dysphoria. A higher level of dysphoria was associated with more frequent use of gait and posture as
cues to get an impression of the affective state of the surveillance targets. Moreover, the degree of perceived arm swinging was found to be related to performance on the Affect Detection task. Better performance of the observers in the estimation of dysphoria level amongst surveillance targets when carrying a firearm was associated with a lower degree of perceived arm swinging in a target.

These results provide more insights into whether the perceived affective state has a potential to be used in order to determine if someone is concealing a firearm. Moreover, the video-footage of surveillance targets when viewed from a street-level perspective used in the present study showed that the viewpoint might have an effect on the judgments of observers, at least in the task when they are required to estimate the affective state of surveillance targets. It has been demonstrated that viewing footage with a full frontal view of the surveillance targets walking towards the camera leads to the observers’ estimation of a targets’ affective state which is congruent with the self-estimated affective state of the targets. The CCTV-level perspective used in study Three seems to trigger the opposite trend.

This finding may well be related to the fact that the surveillance camera positioned at head height provides a better view of the targets’ face than a camera mounted at a height of 2.8 m (i.e., an imitation of the ceiling mounted CCTV cameras), as suggested by the CCTV Operational Requirements manual 2009 (Cohen et al., 2009). Recognition of affective state was better using this head height camera than in the previous studies; this may well be related to the participants using the targets’ face to make judgements about their affective state and their faces were more visible with the head height camera. In addition, this is in accordance with evidence provided by studies aimed to investigate whether the representations mediating the identification of human actions, body postures, or in affect recognition are viewpoint specific (e.g., Daems and Verfaillie, 1999; Coulson, 2004). Coulson (2004) suggested that the frontal view of a person offers more rapid and accurate decoding of emotional state than other views. Since this assumption seems to be supported by the results of the present studies related to the estimation of affect of surveillance targets, it was decided to
continue the research using images of the frontal view of the targets in the subsequent experiments, and to concentrate more on the observers’ visual inspection of different parts of the targets’ frontal view.

A fuller understanding of the visual strategies that observers use during the gun carrier detection task would be gained through experiments that employ eye-tracking. Eye-tracking data would add supplemental information about what areas of interest, defined beforehand, are most informative and visually attended to most or alternatively more frequently ignored by an observer. An eye-tracking system can provide a relatively direct and high-resolution means of capturing information about participants’ visual attention (Vertegaal, 1999). Furthermore, by exploring the visual behaviour of observers allows direct measurement of which parts of the image are relevant for performing the experimental task without relying only on the observers’ consciously reported strategies. Applying this particular method will, therefore, provide information about whether the observers’ perception of particular cues, such as the face or the body of the surveillance targets, is associated with a certain eye-movement pattern, or fixation locations, in order to discern a targets’ affective state. In addition, eye-tracking can provide data which would help determine how observers search the image of a walking person for a targeted object (i.e. firearm).

6.5. Conclusion

The present study provides information about the cues observers use in order to make a decision about the affective state of surveillance targets. Without knowing about the presence of firearms concealed under the clothing of the targets they attributed a negative affect to those surveillance targets with a concealed firearm and viewed from street-level perspective, which was in congruence with the self-report of the surveillance targets themselves. Given that this is the opposite outcome of study Three, such results may suggest that the task of affect recognition might be influenced by the viewpoint. This finding may
not be surprising if the observers were mainly using the recognition of features, such as facial features, which were more visible in a frontal view than in a CCTV-like view. However, the results indicate that according to the answers of observers on the cue-detection questionnaire the gait, posture and arm movements were the main cues which were used in this task. All of these were visible in the CCTV-like view. Consequently further investigation is merited. Additional information about the cues associated with affect detection task and the gun carrier detection task will be gained by monitoring the visual search behaviour of observers which will be discussed more closely in the next chapter.
7.1. Introduction

In the study described in the previous chapter (experiment Four) the observers watched the footage and estimated the emotional state of surveillance targets without knowledge about the presence of a firearm (i.e., Affect Detection task, AD task). The Likert-scale questionnaires were used to obtain information about characteristics of the targets' movement pattern in order to understand how exactly carrying a gun affects the appearance of the carrier to the observer in terms of cues related to the carrier’s emotional state. The questionnaire provided the data on which cues convey information needed to identify the affective state in a firearm and innocuous object carrier.

In the present studies the observers performed the Affect Detection task (experiment Five) and Gun Carrier Detection task (experiment Six). In order to fully understand the strategies which the observers are using during these tasks, their visual behaviour was explored, from which it was inferred which parts of the image are relevant for performing those tasks. According to some researchers (e.g., Nahn et al., 1997, Eastwood, Smilek, and Merikle, 2001; Green et al., 2003) the eye-tracking technique might be able to show whether emotionally arousing information presented in images would have an effect on eye-movement patterns. Applying this particular method may, therefore, also provide necessary information about which eye-movement pattern would accompany the perception of a particular affective state and the discrimination process between a concealed firearm carrier and a carrier of concealed innocuous object.

As mentioned previously (see chapter One), the visual attention distributed between the main cues, such as the face and body of a surveillance target, might be directly related to the task of recognising the affective state of the carrier and thus the carrying of a concealed firearm, since, as research
suggests (Ekman, 1965, Ekman and Friesen, 1967), the perception of non-verbal information about affective state can be obtained through two main channels, facial expressions and body movement. In accordance with previous research (e.g., Montepare et al., 1999; Paterson et al., 2001) the questionnaire from the previous study (experiment Four) consists of questions about the use of such cues as facial expressions, arm swinging and length of stride. These are related to the face, the upper body and the legs of a person respectively. Consequently, it seemed appropriate to divide the images similarly into these separate body areas and to create three areas of interest in the eye-tracking studies, i.e. face, upper body and lower body of the surveillance target, in order to examine the visual attention of the observers.

There is evidence (e.g., Torrabla, Oliva, Castelhano, Henderson, 2006; Brockmole, Castelhano, Henderson, 2006) to support the prediction that visual scanning behaviour would provide valuable additional information about the allocation of visual attention across the scene as eye fixations tend to cluster on some regions of the scene and not at others (i.e. not randomly). The differences in visual scanning behaviour can therefore be attributed to the strategies that observers are implementing while they examine the scene in each task condition. Investigations into eye movement control during scene perception (e.g., Torrabla et al., 2006; Castelhano, Mack, Henderson, 2009) showed that in visual search tasks human observers’ eye fixations largely remain within the scene areas that most likely contain the target object. Moreover, the context information related to visual targets is known to lead to more efficient searches (Brockmole et al., 2006). Thus, the behavioural relevance or local image cues of particular locations in the scene play an important role in controlling the direction of attention (Itti and Koch, 2000). From the research on the evaluation of user interfaces it is known that fixation duration is one of the key indicators of users’ attention for the certain parts and elements of the user interface (Goldberg and Kotval, 1999). Thus, perceptual information processing might be assessed by measuring the mean duration of eye fixations on particular components of an image. Accordingly, the average time the observers spent looking at different areas surrounding bodily features
should provide information about the division of observers’ visual attention in the task of affect / gun-carrier recognition. In particular, this would allow examining which cues were most eye catching, most informative for observers, or most frequently ignored by them. In line with the main purpose of the studies (i.e. examination of most informative cues needed to discriminate between a firearm and non- firearm bearer and to identify their affective state) only mean dwell time in particular areas of interest was considered to be most important metric that could be used in the current studies. Other metrics which are in use in eye measurement research such as pupil-diameter blink rate and saccade speed were expected to be of little value in the present studies. The change in pupil diameter is known to be related mainly to the change in mental workload (e.g., Matthews, Midleton, Gilmartin, Bullimore, 1991; Hilburn, 1996; Blechko, Hilburn, and Zon, 2001). However, the measurement of the mental workload was not included in the current studies. Although the change of pupil diameter (i.e. the dilation) was also found to be associated with perception of strongly arousing negative stimuli presented to test subjects (Partala, Jokiniemi and Surakka, 2000), it was expected that the size of pupil would be much less sensitive in the current settings due to the lower and more subtle intensity of perceived affect. Based on previous research (e.g., Matthews et al., 1991; Stern, Boyer, Schroeder, 1994) it is known that the blink rate and saccade speed or duration are rather useful metrics to investigate the vigilance or visual and mental workload which were not examined in the current study. Furthermore, different studies showed somewhat contradictory results related to the use of eye blink rate as a metric of visual workload. An increased (visual) workload was found to be associated with decline in eye blink rate (Holland and Tarlow, 1972; Hancock, Wulf, Thom, & Fassnacht, 1990; Veltman and Gaillard, 1998) as well as with increased eye blink rate (e.g., Luckiesh, 1947; Mourant, Lakshmanan, and Chantadisai, 1981; Tanaka and Yamaoka, 1993). Consequently, the pattern of eye-movements in the current studies was examined only by estimating the average time the observers spent on looking at different areas surrounding specific bodily features. This is a very common
approach used in many image perception studies where the interest is in which image areas attract attention.

Additionally, in the present study the following question was investigated: how do observers’ visual scanning behaviour and their performance on the Affect Detection task and the Gun Carrier Detection task change dependent on the task demands. Furthermore, by examining the eye movements of observers it was possible to assess how the observers’ awareness about affective behavioural cues, which may accompany a concealed firearm carrying, would affect eye movement behaviour. In order to find an answer to this question one of the two groups of observers was provided with information about these cues included in the experimental instructions, and the performance of both groups on the GCD task and their eye movement patterns were compared.

7.2. Experiment Five

7.2.1. Method

7.2.1.1. Participants

All participants were postgraduate students and members of staff at Loughborough University and were recruited via word of mouth. There were twelve participants, 7 male and 5 female with a mean age of 27.75 ($SD = 2.6$). The participants had no training or experience in any sector of the security industry. All participants were naïve to the purpose of the study.
7.2.1.2. Materials

Apparatus

A Tobii X50 stand-alone eye-tracker with ClearView 2.6.0 software was used to collect the eye movement data. The Tobii system included two monitors connected to the computer allowing participants to view the stimulus on one monitor (i.e., stimulus computer monitor) while the other monitor (the eye-tracking computer monitor) was used by the experimenter to set up the test and watch the eye-tracking quality during the session. For detailed description of the stimulus presentation and the set up of the experiment see Chapter 2, section 2.7.2. Participants sat approximately 60 cm away from this display monitor which provided the stimulus. The AD task’s data collection was automated using another computer, i.e., a questionnaire computer.

Mock CCTV footage

The mock CCTV footage comprised multiple video clips of 11 actors (i.e., surveillance targets), each filmed individually and whilst walking and carrying either a concealed firearm or a concealed innocuous object matched to the firearm for approximate weight and size (i.e., a bottle). Each video clip appeared in the middle of the stimulus computer screen in the form of a rectangle. For the detailed description of the footage used in this experiment see Chapter 2, section 2.7.3.

The duration of each video clip was 4 seconds. Eye movement data over the first two seconds of each video clip were discarded due to the fact that accurate registration of eye movements in each area of interest was only possible for the two last seconds of each video clip because of the size of the target on the monitor screen as they walked towards the camera.

For the purpose of the study three Areas of Interest (AOIs) were defined for each of the 22 video clips: ‘Face’, Upper body (‘Body’), Lower body (‘Legs’) (see Figure 7.1). The three AOIs were designed based on four coordinates indicating a rectangle surrounding the head, the upper body and the lower body
(i.e., legs) of the surveillance target, at its maximum 21 pixels from the edge of the person’s body, bearing in mind the accuracy of 1 degree of visual angle at 60 cm distance from the eyes of observer to the screen of the stimulus computer. These coordinates were noted for each frame in a file of each surveillance target, and the file containing the AOI data was read into the analysis program. The range of the mean height and width of the three AOIs along with the degrees of visual angles at the start point of the walking path and at the end point of the walking path are presented in Figure 7.1.

Figure 7.1. Example of smallest (A) and largest (B) frames from a video sequence, demonstrating eye fixations and three AOIs (i.e., Face, Upper Body, Lower Body) annotated with dimensions in cm and degrees of visual angle for height (H) and width (W) of each AOI.

Affect Detection task
The levels of positive and negative affect were assessed for each surveillance target, both whilst carrying a firearm and whilst carrying the innocuous object, using the Multiple Affect Adjective Check List – Revised, State (MAACL-R; Lubin and Zuckerman, 1999). This was achieved within an experimental design
that was counterbalanced for order of exposure to the firearm and innocuous object.

**Design**

A factorial design with repeated measures was used in this study to vary the type of staged video clips (Innocuous Object versus Gun) as a within subjects factor. Every experimental trial consisted of one video (see Table 7.1), either with a matched object, a concealed bottle (*i.e.*, Innocuous Object trial) or with a concealed firearm (*i.e.*, Gun trial).

Only concealed staged video clips were used in this experiment (for a detailed description of the footage see Darker *et al.*, 2008). There were, in total, 22 video clips with 11 different surveillance targets in two conditions (Innocuous Object and Gun). Each of 11 surveillance targets were shown separately, in one experimental trial; the order of Gun and Innocuous Object trials was counter-balanced across participants.

<table>
<thead>
<tr>
<th></th>
<th>Trial Innocuous Object (Bottle)</th>
<th>Trial Gun (Firearm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concealed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking towards the camera</td>
<td>1 x 1 litre (Total 11 clips)</td>
<td>1 x Glock (Total 11 clips)</td>
</tr>
</tbody>
</table>

*Table 7.1. Overall number of clips used in the study in each kind of trial (Innocuous Object and Gun) with different kind of carrying object (*i.e.*, bottle and firearm), for each surveillance target.*

Dependent variables in this study were: (i) Performance of observers on the Affect Detection task (*i.e.*, mean scores observers gave on scales of MAACL-R questionnaire; differences between the surveillance targets’ and the observers’ scores on MAACL-R); (ii) Mean dwell time (*i.e.*, the mean time in ms spent looking at one bounded area) on each of the three AOIs.
7.2.1.3. Procedure

After arriving, participants read and completed a Loughborough University approved informed consent form. Participants were then instructed to sit close enough to the stimulus computer monitor so that the experimenter could see their eyes on the eye tracking part of the stimulus computer display. The participants were seated 60 cm from the display. Calibration of the eye tracking was completed by asking participants to look at a series of nine dots that appeared on the stimuli computer monitor. Upon completion of calibration, participants were asked to focus their eyes on the white dot in the middle of the stimuli display, and thereafter to view a video clip. They were instructed then to watch the video and after that to make an estimation of the affective state of the person in the video. Therefore, the MAACL-R questionnaire was applied; it was presented on the questionnaire computer. After that the procedure (except the calibration) was repeated. The test session was preceded by a short training session. Following the completion of the experiment the participants were thanked and debriefed.

7.2.2. Results

*Performance on Affect Detection task*

A non-parametric alternative for the paired samples t-test, Wilcoxon Signed Ranks Test was performed in order to investigate the effect of the Type of Object Concealed (Gun versus Innocuous Object) as the factor.
Figure 7.2. The surveillance targets’ affective state, estimated by observers. The surveillance targets’ affective state is represented by the mean scores on the scales of MAACL-R in Innocuous Object (i.e., carrying an bottle) and Gun (i.e., carrying a firearm) conditions.

The results showed that the observed dysphoria in the Gun condition was higher than the observed dysphoria in the Innocuous Object condition (see Figure 7.2), which was congruent with the self-ratings of the surveillance targets. However, this effect was not significant.

Eye movement data

The mean dwell time in each AOI was computed by summing all gaze points enclosed in the AOI and by dividing their duration by the number of frames across all the surveillance targets.

The results of the Friedman test show that there are significant differences in mean dwell time across the three AOIs ($\chi^2 = 18.5$, df = 2, $p<.05$ for the Innocuous Object condition, and $\chi^2 = 18.5$, df = 2, $p<.05$ for the Gun condition. Mean dwell time across all participants on the ‘Face’ was significantly longer than on other AOIs in both the Gun and Innocuous Object conditions.
Mean fixation duration across all participants on the ‘Body’ was significantly longer than on ‘Legs’ in both the Gun and the Innocuous Object conditions (see Figure 7.3 and 7.4 for comparison).

**Figure 7.3.** Mean dwell time within the different AOIs (i.e., ‘Face’, ‘Body’ and ‘Legs’) in Innocuous Object (i.e., carrying an bottle) and Gun (i.e., carrying a firearm) conditions.

The results show that in the Gun condition, the mean dwell time on the area of the face of surveillance targets was slightly longer than in the Innocuous Object
condition. However, this difference in the observers’ eye fixations duration on all AOIs between the Gun and Innocuous Object conditions did not reach statistical significance (see Figure 7.4).

**Relationship between performance on the AD task and the visual cues used by observers in this task**

A Spearman’s non-parametric correlation test was performed in order to investigate the possible relationship between the performance of observers on the AD task and the visual cues they used whilst performing the task.

When the firearm was present a number significant correlations were found. There was a significant positive correlation between the perceived level of depression, hostility, dysphoria and the dwell time on the AOI ‘Legs’ ($r_s = 730, \ n = 12, \ p = .007$; $r_s = .628, \ n = 12, \ p = .029$, and $r_s = .615, \ n = 12, \ p = .033$, respectively). This means that the higher level of perceived depression, hostility and dysphoria was related to a longer looking time at the legs of the surveillance targets.

Moreover, a significant correlation was found between the performance on the AD task (i.e., difference between surveillance targets’ and observers’ scores on MAACL-R) and the mean dwell time within different AOIs. There was a significant negative correlation between the mean dwell time within AOI ‘Legs’, and the observers’ performance on the AD task, for instance on the Hostility scale ($r_s = -.640, \ n = 12, \ p = .025$). This result shows that the more correct recognition of hostility in the surveillance targets when they were carrying a concealed firearm was associated with a longer looking time at the legs of the targets.

Compared to the Gun condition, in the condition when an innocuous object was present, no significant correlations were found.
7.2.3. Discussion

This experiment was designed to examine which parts of the image are relevant for performing the task in which the observers needed to identify the affective state of surveillance targets who were filmed previously whilst carrying a concealed firearm. The observers were not informed about the presence of concealed objects (*i.e.*, firearms and matched innocuous objects). The eye movement registration was used to provide data on which cues convey information needed to identify the affective state in a firearm and non-firearm carrier.

The results of this study show that whilst estimating the affective state of surveillance targets the observers’ attention was directed mainly to the face regions of the targets across all the conditions (gun and innocuous object). The observers tended to look on the upper body (including arms and chest of the targets) longer than on the lower body (*i.e.* legs) of the targets. In addition, the results of this study suggest that the more correct recognition of hostility in surveillance targets when they were carrying a concealed firearm was associated with longer looking time of observers’ at the legs of the targets. Although in the current study the observers watched two different groups of video clips, they were not informed about it, and were therefore not consciously aware of the presence of, concealed firearms and innocuous objects. Consequently, the aim of the next study was to investigate the direct influence of knowledge about the presence of concealed firearms by measuring eye-movements and the performance of observers on the task of discrimination of a carrier of a concealed firearm from a carrier of a concealed innocuous object. Assuming that affect recognition might underlie the detection of a person concealing a firearm, the effect of this was also investigated. One possible way to do this is to include information about which affective cues individuals inherently produce whilst carrying a concealed firearm in the experimental instructions. Consequently, another purpose of the study was to examine the differences in performance on the gun carrier detection task and in visual scanning patterns in two groups of observers, *i.e.* those who received
instructions with affective cues and those who received the same instruction but without affective cues.

7.3. Experiment Six

7.3.1. Method

7.3.1.1. Participants

Twenty four postgraduates and members of staff from Loughborough University (16 male and 8 female) volunteered to participate in the study. Their mean age was 29.79 (SD = 5.5). The participants (except one) had no training or experience in any sector of the security industry. All participants were naïve to the purpose of the study.

7.3.1.2. Materials

Apparatus
The apparatus employed in this study was the same as in the previous study. The data collection related to the performance of observers in the gun detection task was automated using a computer, i.e. a questionnaire computer and controlled by a program developed in-house.

The performance of observers on the GCD task was analysed using Signal Detection Theory. For a detailed description on how the SDT was applied to this experiment see Chapter 2, section 2.4. In the present study there were two conditions: signal present (i.e., the concealed firearm is featured in the clip) and signal absent / noise (i.e., the concealed bottle is featured in the clip).

The measure of sensitivity to firearms in a gun carrier detection task was the parameter $d_a$, which offers a direct index of sensitivity to the signal (i.e.
concealed firearm) when the frequency distributions of signal present and noise are of unequal-variance.

**Design**

A mixed factorial design was used in this study to vary the Type of staged video clips (Innocuous Object versus Gun) as a within subjects factor, the Type of AOI (Face versus Upper Body versus Lower Body) as another within subjects factor; the Type of Instructions (Affective Cues versus No Affective Cues) as a between subjects factor.

Dependent variables in this study were: (i) Performance of observers on Gun Carrier Detection Task (*i.e.*, GCD task), measured in terms of the Sensitivity parameter $d_{a}$ derived from SDT based analysis; (ii) Mean dwell time on each of three AOIs.

**7.3.1.3. Procedure**

The procedure was similar to that of the previous experiment. However in the current experiment, the participants were randomly assigned to two conditions. In one condition (*i.e.*, Affective Cues condition) the experimental instructions started with the following introduction: 'It is known that carrying a firearm might change the affective state and the non-verbal behaviour (*i.e.*, body movement, posture, gait, facial expressions etc.) of the carrier. It is also known that carrying a firearm can cause the negative feelings by the carrier, as for example, anxiety, hostility or depression. As result of it the posture or the way of walking of a gun carrier can appear to observers as anxious, hostile or depressed…' The instructions for the other group of participants did not contain this introduction (*i.e.*, No Affective Cues condition). Both groups of participants were provided with the images including sizes of concealed objects (for the full text of the instruction see Appendix 7).

After the calibration, participants in both the Affective Cues and the No Affective Cues conditions were asked to view a video clip. After each video clip they were instructed to indicate whether they thought the person in the clip was
carrying a concealed firearm. Initially, participants viewed two practice video clips. They then viewed 22 experimental video clips: 11 clips of people carrying concealed firearms and 11 clips of people carrying concealed bottles.

The duration of each video clip was 4 seconds. The recording of eye movements during the last two seconds were used for the analysis of eye movement data. The order of presentation of video clips was pseudo-randomised which allowed an equal spread of the signal detection conditions across the time-course of the experiment. The observers indicated their confidence as to whether a firearm was present or absent on a six point scale: “Definitely no”; “Probably no”; “At a guess, no”; “At a guess, yes”; “Probably yes”; and “Definitely yes”. Following the completion of the experiment the participants were thanked and debriefed.

7.3.2. Results

Performance on firearm carrier detection task

In order to calculate sensitivity to the firearm within a signal detection framework the frequency of responses across the six response categories in the signal absent and signal present conditions were used. Performance was indexed in terms of the parameter $d_a$ which was obtained by means of a maximum likelihood algorithm and based on a Gaussian distribution (RSCORE 5.3.2; Harvey, 2001).

Mean Sensitivity $d_a$ was higher in the group with Affective Cues than in the group with No Affective cues (see Figure 7.5), but these differences between two groups did not reach statistical significance.
Figure 7.5. Sensitivity (da) to a concealed firearm amongst two groups of participants (i.e., who received instructions with description of affective cues and who received instructions without this description)

Eye movements’ data

In order to analyse eye movements data, a three-way mixed ANOVA was conducted on these data using Type of Video Clip (Innocuous Object versus Gun) as a within subjects factor, AOI (Face versus Body versus Legs) as a within subjects factor and the Type of Instruction (Affective Cues versus No Affective Cues) as a between subjects factor.

The results show that for the main significant effect of AOI on mean dwell time \( F (1.30, 28.73) = 16.85, \ p < .001, \ \omega^2 = .66 \), the assumption of sphericity was violated. Mauchly’s test indicated that the assumption of sphericity had been violated for the main effect of AOI, \( \chi^2(2) = 15.92, \ p < .001 \). Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (\( \epsilon = .65 \) for the main effect of AOI).
There was also a significant interaction effect between the AOI and the Type of Instruction, $F(2,44) = 13.82, p < .001, \omega^2 = .62$, a significant interaction effect between the AOI and Type of Video Clip, $F(2,44) = 5.30, p < .05, \omega^2 = .44$, and a significant interaction effect between the AOI, Type of Instruction and Type of Video Clip, $F(2,44) = 6.29, p < .05, \omega^2 = .47$. The main effects of Type of Video Clip and Type of Instruction, as well as the interaction effect between those two factors were not significant, $F(1,22) = .94, p = .34, F(1,22) = .029, p = .87, \text{ and } F(1,22) = .001, p = .98$, respectively.

The results show that the observers’ viewing patterns differed between the three AOIs with respect to received instructions (all differences are significant at $p < .05$), for comparison see Figure 7.6. In both conditions, Innocuous Object and Gun, the observers tended to look longer at the face of the targets than at the upper body or the legs, when they received affective cues. On the other hand, when observers received no affective cues they looked significantly longer at the upper body of targets than on the face or the legs. Thus, it can be concluded that the affective cues included in instructions had an effect on the eye movements of the observers.
Figure 7.6. Comparison of the effect of Type of Instruction (with affective cues versus without affective cues) on mean dwell time of observers within three AOI (i.e., Face, Body and Legs).

Concerning the differences in observers’ eye movement patterns derived from watching different types of video clips, it was found that when surveillance targets were carrying a concealed firearm, there were significant differences
between the two groups of observers (i.e., who received instruction with affective cues and without affective cues) in their mean dwell time within the AOI ‘Face’ and within the AOI ‘Body’ (see Figure 7.7.A).

**Figure 7.7.** Comparison of the effect of Type of Object (A - gun versus B - innocuous object) on mean dwell time of observers within three AOI (i.e., Face, Body and Legs).
The observers who received the affective cues looked longer at the face of targets and less long at the upper body of targets than the observers who received no affective cues. Regarding the mean dwell time spent on the legs of targets, there were no significant differences between two groups of observers.

A similar pattern was found when surveillance targets were carrying a concealed innocuous object (see Figure 7.7.B). The observers who received instructions with affective cues tended to look significantly longer at the face and less long at the legs than those observers who had not received affective cues. Regarding the mean dwell time spent on the upper body of targets there were no significant differences between the two groups of observers.

In addition, a Spearman’s non-parametric correlation test was performed in order to investigate the possible relationship between the performance of observers on the Gun Carrier Detection task (i.e. parameter $d_a$) and the mean dwell time on each AIO whilst performing the task. No significant correlations were found.

### 7.3.3. Discussion

The main task of this study was to examine the eye movements of observers during viewing video footage of walking surveillance targets with either concealed firearms or innocuous objects and performing the task of identification of a firearm bearer. The effect of the description of affective cues included into the instructions on observers’ performance on this task and on their visual scanning behaviour was examined.

There was little effect of the type of the carrying object (i.e., carrying firearm versus carrying an innocuous object) found on the performance of observers on GCD task (i.e., the sensitivity to concealed firearm). Neither was there an effect on the observers’ visual scanning pattern in the two conditions.

As in the first study, the results of the current experiment demonstrate that the face area attracted a significant proportion of eye fixations that resulted in a longer looking time at the face of the surveillance targets during the task of recognition of a concealed firearm carrier. Although this finding in particular
suggests that the face area of surveillance targets might be the most informative for observers in their decision process related to the GCD task, followed by the upper body area and lower body area, the differences in eye movements between the three AOIs were found to be dependent on the type of instruction (\textit{i.e.}, with and without affective cues). When observers were aware about which affective cues may accompany concealed firearm carrying, they tended to look considerably longer at the face than at the upper body or legs of the targets. However, when they received no information on affective cues then the upper body regions received the most visual attention compared to the regions of face or legs of the targets.

### 7.4. Conclusion

The aim of the two present studies was to provide information about the cues which might be reflected in the body language of those carrying firearms and might be apprehended by observers at a subconscious or conscious level, by means of eye tracking registration. Although previously there was an attempt to examine this using questionnaires (see study Four, Chapter 6), there was a need to obtain more reliable data by applying the eye tracking method. The questions asked in the present studies were therefore how visual scanning behaviour of observers would change depending on the task (Affect Detection task and Gun Carrier Detection task) and which cues in particular (\textit{i.e.}, face, upper body or lower body of surveillance targets) are being used by observers in their performance on these tasks.

To do so, in the first study the observers were asked to identify the affective state from human walking movements accompanying the carrying of concealed firearms, without knowing about the presence of concealed objects (\textit{i.e.}, firearms and matched innocuous objects). By arguing that the decision-making process of the observers in their task of affect recognition might be biased by the knowledge of the presence of firearms, the participants in this first study were not informed about the fact that they were watching different
types of video clips. On the contrary, in the second study the only task of the observers was to identify a bearer of a concealed firearm from the same video footage. Besides that, in the second experiment the manipulation with two different task instructions (i.e., by way of including the affect related cues into the instructions for one group of participants) was applied. This experimental design was applied based on the assumption that recognition of affective states may influence the recognition of a bearer of a concealed firearm.

The results of both studies show that the observers’ attention was driven to specific bodily features whilst making an estimation of the affective state of surveillance targets or whilst performing the task of recognition of a firearm bearer. In the first study this was mainly manifested in a consistent interest in the face regions across all the conditions (gun versus innocuous object). In the first study, observers looked in general also at the upper body (including arms and chest of the targets) significantly longer than on the lower body, i.e. legs of the targets. Furthermore, the described pattern did not differ between the two conditions, i.e. when surveillance targets were carrying a firearm and when they were carrying an innocuous object.

Although the eye movements of observers were not made at random in the current Affect Detection task, but clustered consistently around the face regions of surveillance targets, and the attention to this region was most prominent throughout the whole experiment, the results also suggest that when the targets were carrying a concealed firearm, their legs area, may have been used by observers as a cue to estimate the targets’ affective state. When the target was carrying a firearm the duration of the observers’ eye-fixations on the lower body (i.e., legs) of targets was related to the estimated negative affect whilst identification of affective state. Higher level of perceived dysphoria, hostility and depression was related to longer fixation time on the legs of targets. Besides that, the better performance of the observers on the Affect Detection task, and in particular in their estimations of hostility, was related to the longer time participants spent looking at the legs of surveillance targets when they were carrying a concealed firearm. In other words, more accurate estimation of negative affect in this experiment was related to longer times the
observers spent looking at the lower body of the surveillance targets. The debriefing interview also showed that according to some participants they looked at the legs because they noticed differences in speed or stride lengths of the walking targets. The attention to targets’ legs which might be, indeed, related to speed of walking, implies that cues such as speed of walking or the length of strides may be informative in the task of affect recognition of carriers of concealed firearms.

Similarly to the results of the first study, the second experiment showed that the visual scanning pattern of observers did not significantly differ between the two conditions, *i.e.* when surveillance targets were carrying a firearm and when they were carrying an innocuous object. Accordingly, the observers could not discriminate significantly between a bearer of a concealed firearm and a bearer of a concealed innocuous object. However, there was a tendency toward better performance on this task (*i.e.*, higher sensitivity) for the observers who were informed about the affective response on firearm carrying represented in a change of body language.

Moreover, this study provides further evidence that in both conditions observers’ viewing patterns were influenced by the upper body of the surveillance targets, and particularly by their face. The effect of the instructions, with and without information about affective non-verbal cues associated with firearm carrying, was also evaluated. The instructional conditions are known to be very important for perceptive processes, as it is directly related to attention (*i.e.* the way an individual makes selection within various categories of information), and the insertion of certain words in the intentional instructions can favour the participants’ performance (Rodrigues and Marques, 2006). The application of different types of instruction in the current study was able to induce varying emphasis on the use of affective state as a cue in the task of detection concealed firearm carrying. With the emphasis explicitly put on possible affective cues associated with concealed firearm carrying, the observers in the current studies were enabled to use this information in order to recognize a bearer of a concealed firearm. When no specific instruction was given, the detection of a person with a concealed firearm was expected to be
based on more salient, physical signs (e.g., hand location or a bulge somewhere in the clothing of the target) as shown the study of Darker et al. (2009) than on any other, more subtle cues such as the change in non-verbal behaviour. Indeed, the different types of instructions seemed to have contributed to the differences in observers’ visual scanning behaviour. The manipulation with two types of instruction (i.e., with and without affective cues) showed differences in eye movements between the three AOIs depending on the type of instruction. When observers received information about affective cues they tended to look longer at the face than at the upper body or legs of the targets. In contrast, when instructions contained no information about affective cues then the upper body regions received the majority of eye fixations compared to the regions of face or legs of the targets.

Taken together, the finding of both studies demonstrated that when recognition of affect was included in the task, the faces of surveillance targets seemed to attract most of the observers’ attention, followed by upper body and then the legs of the targets. In accordance with previous research (e.g., Ekman, 1965; Ekman and Friesen, 1967) concerning the informative value of the human face in recognition of the type of emotion, the findings from the first study (i.e. experiment Five) showed that the observers’ attention was driven mainly to the face of a surveillance target when the observers were asked to identify a target’s affective state. Similarly, experiment Six showed that awareness about the possible effects of gun carrying on the affective non-verbal behaviour of the bearer could cause a shift of visual attention to the face of the targets. However, when the task of the observers was to identify the bearer of a concealed firearm without knowledge about related affective behavioural cues, a shift of visual attention to the upper body of the surveillance targets was perceived. In this case, the upper body, including arms and chest, of surveillance targets and to a lesser extent their face and legs seems to be essential for the identification of a surveillance target with a concealed firearm.

Regardless of the noticeably predisposition of observers to look at the face of the targets when the affective state was considered while making a
decision about the presence of concealed firearm, there was no further relationship found between the time observers spent looking at the face of targets and the performance on this task. Although there was no clear indication that prominent attention to the face or facial expressions was likely to enhance the performance of observers, there was a tendency for better performance on the task of detecting a firearm carrier (i.e., higher sensitivity) when the instructions with information about the affect-related cues were provided. Therefore, it can not to be concluded with certainty that use of the face as a cue was sufficient for the better performance on this task. Nevertheless, the results of the current eye-tracking studies may suggest that the face of a surveillance target might be fixated automatically and frequently in order to obtain information about their affective state, despite the fact that the features of the face were insufficiently visible in the video footage, and perhaps it could not reveal subtle changes in affective state. A potential explanation for these finding might be that faces in general tend to capture attention more rapidly than other objects and they may be detected and perceived even outside of attention (Suzuki and Cavanagh, 1995; Mack and Rock, 1998). Furthermore, other evidence from the literature suggests that faces seem to have a special status, and the role of facial expressions in providing information for people making judgments about another person's emotional state has been emphasized (Ekman, 1965; Ekman and Friesen, 1967; Ekman, 2003). Faces in general and specific facial expressions can provide affective information which is of great importance for human social interaction (Mack and Rock, 1998; Ekman, 2003). Moreover, evidence implies that detection and orienting of attention to faces might be facilitated by emotional expressions, as some emotion-related information can be perceived implicitly outside conscious awareness (Esteves, Dimberg, Öhman, 1994; Eastwood et al., 2001; Vuilleumier and Schwartz, 2001).

The fact that the face of the surveillance target in the current study may have been seen by human observers as a possible source of affect-related information, followed by upper body and then the legs of the surveillance targets, needs to be taken into account when considering an affect related
cueing method as a method useful in terms of recognizing (concealed) firearm carrying. The explicit experimental instructions that addressed the affective behavioural cues associated with firearm carrying, seemed to have contributed to the observers’ ability to recognize the bearer of concealed firearm. In consequence, it can be assumed that the performance of human observers in a similar task undertaking in a natural setting would become more effective when explicit instructions to consider affective cues associated with firearm carrying that can be readable from non-verbal behaviour in general and from facial expressions in particular, are provided. This should be especially useful in the situations when other ‘suspicious’ visual cues, such as for example gun shaped bulge in the clothing, are present and visible to an observer. In terms of practical applications it may also imply that high quality CCTV footage is required in order to achieve this. It is known that poor-quality CCTV footage is problematic in such tasks as, for example, the establishing someone’s identity from perception of the face (Henderson, Bruce and Burton, 2001; Keval and Sasse, 2008). As suggested by previous research (e.g., Keval and Sasse, 2008-1; Keval and Sasse, 2008-2), video footage with a high level of resolution (e.g., 352 x 288) and a high frame rate would be beneficial in terms of correct real-life detection of crime events in general (i.e., 8 frames per second and above) and in terms of gathering more detailed and more reliable information from the face of surveillance targets in particular (i.e., 15 frames per second and above).
CHAPTER 8: THESIS DISCUSSION and CONCLUSION

8.1. Overall discussion

This chapter summarises the findings of the present research. The contribution to knowledge, possible application of the research and further work to be undertaken is highlighted as are limitations of the studies. Finally an overall conclusion to the body of research is drawn.

As part of the MEDUSA project, which was concerned with developing software to automatically detect individuals carrying weapons, it was important to investigate how human observers went about identifying potential gun carriers and to determine what cues they consciously or sub-consciously attended to when they were doing this. In particular if it can be shown that non-verbal cues are utilised successfully by human observers then it is possible to develop software to emulate this.

This thesis therefore presents research aimed at investigating whether the recognising of an individual with a concealed firearm through CCTV imagery might be based on an understanding of the human affective state as derived from non-verbal behavioural cues. In other words, the main research question to be answered was: what is the potential value of affective behavioural cues which may accompany concealed firearm carrying. The skills to detect people engaged in unlawful activities (e.g., carriers of illegal firearms) through observation of their affective state and non-verbal behaviour have been recognized to be important in the work of police officers (Burns, 2006; Johnson, 2007). So far, insufficient applied research has been done with respect to the work of CCTV operators, even though their primary task is to detect human mal-intent which may include detection and responding to incidents which involve the illegal use of firearms. The fact that firearms are more often being
carried concealed under the clothing than in the open view makes the recognition of the firearm bearer a very difficult task. It is possibly highly dependent on such factors as individual abilities of human surveillance operators, appropriate training and experience. It is therefore hypothesised that the recognition of offenders with a concealed firearm by a CCTV operator might include the recognition of their affective state based on the perception of non-verbal behaviour that is accessible from the CCTV images, along with a perception of more obvious visual cues (e.g., suspicious gun shaped bulge in the clothing or position of the arms etc.). Additionally, it was assumed that the experience in surveillance and innate abilities to decode affective non-verbal cues would have an effect on the performance on this task. For this purpose evidence from experimental psychology and applied research were reviewed, and a number of experimental studies have been performed.

Evidence from literature showed that carrying an illegal firearm can be associated with a range of emotions, and that even pictures of firearms can automatically prime aggressive thoughts and promote aggressive behaviour (Anderson et al., 1998; Klinesmith et al., 2006). Additionally, through reviewing the literature of empirical approaches to the study of human affect recognition, the thesis has discussed the human ability to detect and understand (i.e., decode) affective states of others and the role of facial expressions and human movements in this respect. The skills relating to the understanding of non-verbal communication include the ability to decode or interpret non-verbal cues transmitted by others (Zuckerman et al., 1978). An overview of the existing research showed how emotional state is reflected in non-verbal behavioural cues, such as facial expressions and human motion (e.g., gestures, arm movements, gait etc.), and how this is decoded by other people. The results of research studies suggest that the emotions of other humans can be perceived and recognised even from simple representations of human movement as for example in a point-light stimulus display (Johansson, 1973; Pollick et al., 2001). Moreover, studies involving functional neuroimaging show that perceived biological movement is processed by the brains of human and non-human primates in a way that supports the automatic inference of the intentions of
others (Blakemore and Decety, 2001; Kohler et al., 2002). Considerable evidence indicates that faces may represent a class of their own (Purcell and Stewart, 1986) and can be processed by the brain as a separate entity (Kanwisher, McDermott, and Chun, 1997).

Referring back to the evidence that certain affective processes amongst offenders were found to be associated with the act of committing a crime and with using a firearm, the assumption that affective states elicited by carrying a firearm might be reflected in changes in the individual’s body language has been discussed. The emphasis of the current research was therefore to look at whether, when attempting to detect the carrier of a concealed firearm, the human observers would respond to the change in non-verbal behaviour of the bearers by attributing different affective states to the surveillance targets.

Furthermore, the research investigated whether accuracy of their judgments (i.e., the degree to which this corresponds to the self-judgments by the targets themselves) might vary and might be related to the observers’ innate affect decoding abilities and certain visual cues associated with firearm carrying. Because the cognitive processes involved in the detection of firearm bearers and in the recognition of the affective state of those subjects were believed to interfere with each other, two tasks were developed and used with the purpose of investigating performance on both tasks separately. Based on this, experimental studies were designed in order to find evidence that carrying a firearm might be associated with negative affect, and to examine how this affective state would be perceived by human observers dependent on their experience with surveillance and non-verbal sensitivity. Moreover, the studies that were considered would provide information about which visual cues the observers would use in both the affect detection and gun carrier detection tasks.

By using generated video images of individuals walking whilst carrying, concealed on their persons, either firearms or matched innocuous objects, the influence of gun carrying on the affective state of the carrier was assessed. It was found that firearm carrying was associated with increased negative affect represented by dysphoria, i.e. a construct created by combining the scores for
anxiety, hostility, and depression scales derived from the self-report measure of affective state. In the subsequent experiment (i.e., study Two) the ability of CCTV operators and lay people to detect whether or not an individual, captured on CCTV, is carrying a concealed firearm was investigated. Firstly, this study examined the potential influence of expertise on the performance on this task, assuming that the CCTV operators, due to their training and work experience, would be able to identify a concealed gun carrier more accurately than people without any training or experience. Secondly, the study examined whether performance in detecting someone carrying a concealed gun is related to the surveillance target’s affective states and to the observers’ recognition abilities of these affective states derived from behavioural cues which individuals inherently produce whilst carrying a concealed firearm. The results showed that the observers’ performance was found to be less influenced by training than expected. Professionals and lay people performed on the gun carrier detection task with relatively the same level of performance, although there was a tendency for CCTV operators to demonstrate higher levels of sensitivity to firearms. One of the possible explanations might be that the experiment was underpowered that suggests that a significant difference might be found using a larger sample in future studies. Furthermore, the task that the participants were required to do in this study could be too demanding in terms of perceptual and cognitive load. The visual and cognitive fatigue that could occur during this task may have influenced the performance of participants. Therefore, in order to reduce the effects of fatigue in the following studies fewer video clips were used.

Another two important findings were that observers’ decisions concerning the detection of concealed firearms appeared to have been associated with anxiety and sensation seeking experienced by the surveillance target. The size of the influence of carrying a firearm on a surveillance target’s anxiety level was found to be related to the number of times that the individual was deemed to be carrying a firearm, but not to the observer’s firearm detection sensitivity. The size of the influence of carrying a firearm on a surveillance target’s sensation seeking level was found to be related to sensitivity in firearm
detection by the observer. This suggests that anxiety and sensation seeking in a surveillance target might be the most visually apparent of the five types of affect considered. These results imply that the decision process regarding whether or not a surveillance target was carrying a concealed firearm may have involved an interpretation of the surveillance target’s affective state. It remains to determine what these visible, behavioural correlates of affective state might be. Additionally, the decoding abilities of participants, as measured by the PONS test in this study, were little related to their performance on the gun carrier detection task. A test of body language decoding ability did not show that body language reading skills of observers were related to their ability to detect a concealed firearm. Another attempt to investigate this relationship was made in the subsequent study. Taken together, these initial results provide some insight into the potential for using a surveillance target’s body language related to their affective state to determine whether they are concealing a weapon, and whether or not such an ability can be acquired through surveillance training and experience.

The next two studies investigated whether observers are able to perceive differences in the affective states of people who are and who are not carrying concealed firearms, as judged by monitoring staged CCTV footage. Similarly to the previous experiment, the participants viewed mock CCTV clips of individuals walking. Half of the clips featured bearers of concealed firearm; the other half featured bearers of concealed, innocuous object. In the first study (i.e., study Three) the observers were presented with footage made from a CCTV-like perspective. They were asked to identify the affective state which they think individuals in the videos might experience, without informing the observers about the presence of concealed objects carried by those individuals. In study Four the observers were presented with the images of the same individuals made from a different point of view, i.e. street-level perspective, in order to investigate possible effect of viewpoint dependence on the performance. The results of study Three showed that observers were able to differentiate between the two clip types by attributing different affective states to the individuals featured in the clips. However, contrary to expectations, their
estimation of affective state was dissimilar and in the opposite direction compared to the affective state reported by those featured in the clips, *i.e.* they attributed higher levels of negative affect to surveillance targets when they were carrying an innocuous object and higher levels of positive affect when the targets were carrying a firearm. The results obtained in study Four showed that the observers could discriminate between the targets’ affective states, but in contrary to the previous study, the estimation of negative affect was correct. This result then provides a possible explanation for the findings in the previous study. The observers’ recognition of affective state of surveillance targets was enhanced, possibly due to presenting observers with a frontal view of the surveillance targets (*i.e.*, street-level perspective) instead of a CCTV-like perspective. This finding is consistent with previous research which showed that different angles of viewing faces and body movements result in a different degree of recognition of certain characteristics of the observed subject (*e.g.*, Troje et al., 2005; Troje and Bulthoff, 1996). Moreover, it is believed that the attribution of an affective state to a body posture might become easier when an observer perceives a posture facing the observer compared to other views due to little occlusion of one body part by another. This superiority of frontal view in perception of affect leads to more consensual attributions of affective states to postures presented to observers in the frontal view (Coulson, 2004). In the current research the improved performance of observers on the affect detection task may mean that the street-like perspective might provide observers with more accessible cues to be able to discriminate between affective states.

In addition, the test of affect decoding abilities was performed in study Three in order to examine the relationship between individual differences in sensitivity to non-verbal cues of observers and their performance on the affect detection task. Similar to the results of the study Two, the results confirmed that observers’ non-verbal sensitivity measured by the PONS test and their ability to recognize the affective state of carriers showed little relationship. It may be argued that regarding the relationship between measurement of sensitivity to non-verbal cues and the performance on both affect detection and gun carrier detection tasks, the results of the present research demonstrated little evidence
for the assumption that the decoding ability of observers, measured by the PONS test, can predict the ability to either identify a firearm bearer or to recognize their affective state through the present mock CCTV footage. The lack of expected relationship between measurements of observers’ decoding abilities and their performance on both tasks in the current research can be due to gender differences in non-verbal expression of emotions presented in the PONS test and in the staged video clips. Literature indicates that communication of non-verbal expressions by men and women is different, and for this reason it can be interpreted differently by observers (DePaulo, 1992). The fact that measuring sensitivity to non-verbal cues has been done by applying a test in which only one, female, encoder of non-verbal communication was used, while all actors in the staged clips were male, may have influenced the way they expressed an emotional state, and thus had a different effect on the observers' performance on the PONS test and on the mood detection task. As a result, it may be concluded that the PONS test may not capture the same aspects of reading body language as those used in the detection of a concealed firearm bearers or their affective states, and also it may not be sensitive enough for this kind of task.

Additionally, the visual cues used by observers to judge the affective states of surveillance targets were examined in the present research by applying a questionnaire and by using eye-tracking. The observers' answers on the questionnaire, which was developed for the purpose of study Four, showed that a higher level of negative affect (e.g. dysphoria) was associated with more frequent use of gait and posture as cues in order to get an impression of the affective state of the targets. Another visual cue, the degree of perceived arm swinging, was found to be related to the recognition of affect in targets. A lower degree of perceived arm swinging in a target was associated with the better recognition of the dysphoria level of targets.

Another two studies (i.e., studies Five and Six) that implemented eye movement recording were conducted in order to further investigate the strategies which observers apply whilst performing affect detection and gun carrier detection tasks. The visual scanning behaviour of observers was
examined during performance on both tasks. Research on the topic of scene perception and related visual scanning behaviour shows that the allocation of visual attention across the scene is not a random process as human eye fixations cluster on some regions of the scene and not at another (Brockmole, et al., 2006), and that in the visual search tasks human observers’ eye fixations largely remain within the scene areas that most likely contain the target object (Castelhano, et al., 2009). Thus, it is expected that the visual scanning behaviour would provide valuable additional information about a target search within a scene. In relation to the current research it was therefore assumed that eye-movement data could provide necessary information about to which parts of a target the observer would attend that would accompany the discrimination process between a concealed firearm carrier and a carrier of concealed innocuous object, based on the perception of the carriers’ affective state.

Study Five examined which parts of the image were relevant when the observers needed to identify the affective state of surveillance targets without knowing about the presence of concealed firearms. The results of this study show that the observers tended to look at the upper body (including arms and chest of the targets) longer than at the lower body (i.e. legs) of the targets. However, the face regions of the targets attracted the most attention of the observers across all the conditions (i.e. when targets were concealing a gun and when they were concealing an innocuous object).

Study Six was designed to examine which parts of the image would attract the observers’ visual attention while performing the task in which they were asked to recognize a bearer of concealed firearms. Besides that, the effect of the instructions, with and without information about affective non-verbal cues associated with firearm carrying, was evaluated. The eye movement registration in both studies provided data on which cues convey information needed to perform the two tasks. The results of both studies showed that the observers’ attention was driven to specific bodily features whilst performing the tasks. In a similar fashion to study Five, the results of study Six showed that there was no significant differences found in the scanning pattern of observers when they were watching images with concealed
firearms and images with concealed innocuous objects. In both conditions, the attention of observers was driven to the face, followed by the upper body and the legs of surveillance targets. As mentioned earlier, the face of a surveillance target attracted the most attention from observers when they were performing the task of affect recognition in study Five. The same effect was perceivable when observers needed to detect a concealed firearm carrier in study Six and they were aware about possible effects of gun carrying on the affective state and non-verbal behaviour of the carrier. In other words there was a clear predisposition of observers to look at the face of the targets when the affective state was considered while making a decision about the presence of a possible concealed firearm. On the contrary, the upper body, including arms and chest of surveillance targets, and to a lesser extent their face and legs seem to be important areas for the identification of a surveillance target with a concealed object when observers do not know about related affective behavioural cues. Taken together, the findings of both studies lead to the conclusion that the face of a surveillance target attracts the most attention from observers when they need to identify a target’s affective state, or when they are aware about affective cues related to gun carrying. This is consistent with the idea that the face may need to be fixated more often to obtain information about an affective state as it is more informative about the type of emotion (whether the person feels angry, afraid, sad, etc.) compared to the body which is believed to be more informative about the intensity of an emotion (Ekman, 1965; Ekman and Friesen, 1967). Although the results did not show a clear relationship between the time observers spent looking at the face of targets and the performance on the task of concealed gun carrying detection, there was a tendency for better performance on the task of detecting a firearm carrier (i.e., higher sensitivity to firearms) when observers were provided with information about the affect-related cues. The explicit instructions that addressed the affective behavioural cues associated with firearm carrying seemed to have contributed to the enhanced ability of observers to recognize the bearer of a firearm concealed under the clothing. Thus, there is an indication that the performance of surveillance operators on similar tasks would become more effective when the
instructions about possible affective cues associated with firearm carrying, readable from facial and bodily expressions, are provided. Although it cannot to be concluded with certainty that the attention to the targets’ face in particular was sufficient for the significant improvement of the observers’ performance, the results of the current eye-tracking studies may suggest that the face of a surveillance target might be fixated automatically and frequently in order to obtain information about their affective state. Moreover, the visual attention of observers was directed to the targets’ face despite the fact that the features of the face were insufficiently visible in the video footage to be able to reveal all subtle changes in affective state. Previous research showed that the use of poor-quality CCTV footage decreases significantly the possibility to identify an individual from observation of the face of this person (Henderson, et al., 2001; Keval and Sasse, 2008). Furthermore, the existing contemporary requirements for placement of CCTV cameras (e.g. CCTV Operational Requirements manual 2009, Cohen et al., 2009) suggest that surveillance cameras that are placed at head height provide a full view of the suspect’s face, which would facilitate suspect identification. The present research demonstrated that there is an indication that frontal view of surveillance targets also allows a more precise recognition of their affective state. Besides that, the prominent attention to a face of surveillance targets when affective cues are involved in the task, may suggest that high quality CCTV footage would be beneficial in order to make a target’s face a more reliable source of affect-related information. Further possible applications of the research in this respect will be discussed in the next section.

8.2. Contribution to knowledge / Application of the research

The present research demonstrated that the detection of concealed firearm carrying through CCTV imagery can be linked to the affective state experienced by an individual whilst carrying a concealed firearm. However, it also showed that the recognition of a person who carries a concealed firearm could not be
performed by CCTV operators with a necessary level of precision. It is known that CCTV-mediated surveillance is a complex and labour-intensive activity. The majority of the operator's working time can be spent viewing images in which there is no suspect activity which leads to operator fatigue and inevitably results in errors (Freer, Beggs, Fernandez-Canque, Chevrier, Goryashkot, 1995). The development of an automated visual surveillance system is one of the possible methods for improving the efficacy of CCTV surveillance. Such a system might assist CCTV operators in the prediction of unlawful human behaviour in general, and might also assist in the detection of those carrying concealed weapons such as firearms. There has already been automatic visual analysis technologies developed which allow CCTV operators to be alerted about the possible presence of unusual, suspicious human behaviours, for example through an analysis of human movement patterns (Haritaoglu, Harwood, and Davis, 1998; Hampapur, Brown, Connell, Pankanti, Senior, and Tian, 2003). However, many automatic visual surveillance systems have been shown to perform inadequately because of the unpredictable nature of human behaviour (Goneid, el Kaliouby, 2002) and intentions that is to a large degree dependent on changeable emotional state experienced by individuals. Greater accuracy might therefore be achieved through the integration of the human affect recognition component into the development of automated surveillance techniques. The present research demonstrates that such as a computer-based surveillance system which is aimed to assist CCTV operators in their attempt to detect human mal-intent which includes carrying concealed firearms could benefit from an ability to interpret human affect because it would make the system work in a more similar way to the way the human operators work.

Another application of the current study relates to the fact that humans differ in their individual cognitive and perceptual skills and they need to be trained to perform on a task with as equal level of performance as possible. People can also be expected to infer different meanings for behaviour when it comes to interpreting scenes on CCTV monitors. These differences might be found in their perception and interpretations of affective cues derived from human behaviour. The present research has identified several cues that were
associated with the performance of human observers on affect detection and gun carrier detection tasks. The results show that perceived negative affect that has been known to be associated with firearm carrying was related to observers’ attention to gait, posture and arm movements of surveillance targets. Furthermore, eye movement registration added other valuable information about the way observers were using visual cues, such as the face, upper body and legs of the targets. Perceptual functioning in general is known to have the character of skills. By training, a perceptual sensitivity to human movements can be made sharper, and the skilled observers can be expected to be able to obtain more precise information from human movements (Runeson, 1985). Consequently, it can be assumed that another way to continue to improve human factors issues in CCTV-mediated surveillance is to give more attention to a training program that will address the issue of affect perception from human movements along with cognitive and visual abilities in CCTV operators. The knowledge about the behavioural affective cues accompanied concealed firearm carrying obtained in the present research might be used in further research on developing appropriate training for surveillance operators. However, a number of methodological issues and limitations that could have influenced the outcomes of the current research should be discussed here.

8.3. Methodological issues / Limitations

A number of methodological issues that will be discussed in this section are to point out limitations of the current study, in order to avoid these in future research. As mentioned earlier (see chapter Two) the staged CCTV video clips of walking people carrying real firearms and innocuous, matched objects were used in the current studies due to three reasons. Firstly, the real-life footage of gun crime is extremely rare and difficult to get hold of. Secondly, there was a restricted time period for the collection of the requested amount of CCTV footage. Bearing in mind that the use of mock CCTV footage could be
questioned as it does not hold a very high ecological validity compared to real CCTV footage, the creation of mock CCTV footage was also done deliberately, in order to control a number of experimental variables that are difficult to control in the natural settings (e.g., experimental conditions such as environmental lighting, the positioning of the cameras time of the day, clothing of surveillance targets, overall video image quality etc.), which was the third reason of using mock CCTV footage instead of footage recorded at CCTV control rooms. In this respect, the performance on both the affect detection and gun carrier detection tasks could be measured objectively. To ensure that created footage would realistically represent real-life CCTV imagery and would approach acceptable ecological validity, the footage needed to mimic the live real-time detection of individuals carrying firearms on the person. Therefore, the recording of surveillance targets was performed only after a careful review of real CCTV footage that showed people walking along the street with illegal firearms.

The present research was mainly concerned with the topic of concealed firearm recognition as performed by human observers without expertise in surveillance. In the first study (see chapter Four) the abilities of CCTV operators and lay people to detect concealed firearms in video clips were assessed. The study showed that although it was hypothesised that CCTV operators would demonstrate significantly better performance in the detection of concealed firearms than individuals without any prior experience in surveillance, the task of detection of concealed firearm carrying was found to be equally difficult for both trained and lay people. This might suggest that the skills involved in surveillance activities at a professional level, acquired through training and experience, did not significantly contribute to the performance on this particular task. Comparable results were obtained by other studies (e.g., Troscianko et al., 2004) that showed that CCTV operators and lay people are able to achieve the same level of mal-intent detection via CCTV. For this reason only students were used in the subsequent studies of the current research. However, future studies that would address the effect of training in decoding of affective non-verbal behaviour in similar tasks should rely more on
data obtained from participation of CCTV operators. This would explain more about what such training can add to their already existing work experience.

The fact that the footage was generated in an artificial experimental environment could have some effect on the behaviour of the people who served as surveillance targets in this study. According to Vrij, Edward and Bull (2001) it can be questioned whether human behaviour would be the same in the laboratory, which is a low stress and low risk environment, and in the real-life situation when people are dealing with an official police interrogation or contact on the street. Although it has been assumed that the surveillance targets in the current research showed spontaneous (i.e., not posed) non-verbal behaviour, it remains unknown how the artificialness of the testing situation and tasks was reflected in their non-verbal (experienced and displayed to observers) behaviour whilst they were walking with a firearm in front of the cameras. According to studies of spontaneous expressiveness, an involuntary non-verbal behaviour can convey reliable information about people’s experiences (Edelmann and Hampson, 1981; Ekman, 2003). However, not spontaneous, posed affective state is more difficult to produce and depends on the personal encoding abilities of an individual (DePaulo, 1992). How spontaneous expressiveness of participants whilst being filming and carrying a firearm really was, may be questioned. One could therefore expect that there might be some incongruity in the experienced state by surveillance targets and the appearance of this state to observers in the form of non-verbal behavioural cues.

The following aspects should receive more attention whilst creating idealized CCTV footage for the purpose of experimental sessions. The research presented in this thesis did not examine the possible influence of gender, racial and cultural differences as the footage employed mainly Caucasian male subjects. From the research about face identification from CCTV imagery (e.g., Keval and Sasse, 2008) it is known that human observers consider female faces easier to identify than male faces. Besides that, the ethnicity of a human target was found to influence the degree of difficulty in face identification through CCTV imagery, based on subjective responses and
the task performance of observers (e.g., White-Caucasians faces were reported as the easiest to identify and Afro-Caribbean faces were the most difficult to identify). It can be assumed that these findings may also have implication for the interpretation of affect from facial and bodily expressions. Indeed, in relation to the topic of cross-cultural recognition of emotions it has been argued (e.g., Russell, 1994; Elfenbein and Ambady, 2002) that the human language of emotional expression may have aspects that differ in the style of expression and interpretation between the cultures and may lose some of their meaning across cultural boundaries.

Moreover, based on research on deceptive behavioural cues (e.g., Winkel and Vrij, 1990; Vrij, Dragt and Koppelaar, 1992) Johnson (2007) emphasizes the possible influence of race and ethnicity on the baseline at which suspicious non-verbal behaviours are being displayed. The results of this research suggest that a person’s non-verbal behaviours vary significantly between different ethnic groups (e.g., Afro-Americans versus Caucasians). The research on deception and suspiciousness in human behaviour (e.g., Vrij, Akehurst and Morris, 1997) shows that there is a large variation between individuals in how frequently they normally display certain non-verbal behaviours. It suggests that in the interpretation of results in the future studies one should be aware that without knowledge of each person’s baseline behaviour it might be difficult to correctly identify whether or not they have increased or decreased non-verbal behaviours. Taken together, these studies suggest that such factors as gender, possible cultural variations in the display and decoding of human affect and person’s baseline behaviour should be taken into the account when considering the future research on similar topics.

8.4. Future research

The findings have demonstrated that the affective state experienced by an individual whilst carrying a concealed firearm can be linked to the detection of a concealed firearm on that person when they are observed via CCTV. Future
research should address whether this phenomenon could be honed to support the reliable detection of subtle cues that might indicate that a surveillance target is carrying a concealed firearm.

At the moment, further work is also needed to clarify how innate human abilities to decode and interpret affective information derived from non-verbal behaviour will affect recognition of concealed firearm carrying. The results of the present research did not show enough support for the assumption that this ability measured by the PONS test correlate with the ability to recognize a concealed firearm carrier. Perhaps another approach should be considered. Meiran et al. (1994) were addressing a question of what are the process differences between efficient and inefficient nonverbal decoders by generated hypotheses which are consistent with the assumption that nonverbal decoding tests measure sensitivity to non-verbal cues. According to the first hypothesis, "ignorance", efficient decoders, compared to inefficient decoders, use a larger repertoire of nonverbal cues. The second hypothesis "overall sensitivity" states that all subjects are sensitive to the same set of cues but efficient decoders are generally more sensitive to the cues than inefficient decoders. Based on these hypotheses it seems useful not to use any existent tests of non-verbal sensitivity but instead, to identify the potential cues that would be used only in the task of identification of concealed firearm bearer, and to analyse the differences in sensitivity to those cues between efficient and inefficient decoders, similar to the way it was done in the study of Meiran et al. (1994). Future research on the relevant topic is needed to investigate the possibility to compare efficient and inefficient decoders in the task of detection of concealed firearm bearer.

Another topic worthy of further research concerns the influence of viewpoint on the detection of (concealed) firearm carrying through CCTV imagery. The subject of the optimal camera placement is vitally important in real-world surveillance which includes motion recognition, as this can maximize the observability of the motions taking place (Bodor et al., 2007). Furthermore, a positioning of a camera or a set of cameras in order to effectively observe the area of interest is a challenging problem also because an optimal camera
placement seems to vary depending on different task requirements, whether it is subject tracking, activity classification or gesture recognition (Bodor et al., 2007). The issue of proper camera placement for the purpose of optimizing the camera’s sensor’s ability to capture information about a desired environment or task has been investigated in a number of studies (Fiore et al., 2008). In particular, the topic of enhancing visibility and depth of field (e.g., Tarabanis, Tsai, Kaul, 1996), and the problem of maximizing camera coverage of an area using different cameras’ field of view (e.g., O’Rourke, 1987) have been studied considerably. These studies suggest that camera placement is an important issue to consider in the research related to detection of concealed firearm carrying through CCTV. Although it has not been investigated properly yet, the present research has shown that there is an indication that the recognition of affective state related to firearm carrying might be viewpoint dependent. In the street-level perspective the surveillance targets were correctly judged to have higher levels of dysphoria whilst concealing a firearm than whilst concealing an innocuous object, but from the CCTV-level perspective the opposite trend was observed. From a number of investigations it is also known that the viewpoint can affect the recognition of human actions, gender and affective states (e.g., Verfaillie, 1993; Mather and Murdoch, 1994; Coulson, 2004). According to experimental evidence, frontal views in particular are known to lead to more consensual attributions of affective states to presented postures (Coulson, 2004). Future studies could investigate this research area related to detection of mal-intent in general and the firearm carrying in particular more closely, for example by creating and comparing CCTV images of surveillance targets with frontal and side view.

8.5. Overall conclusion

The main purpose of the research was to investigate whether the recognising of an individual with a concealed firearm through CCTV imagery might be based on
an understanding of human affective state derived from non-verbal behavioural cues.

Evidence reviewed here also suggests that there is no effect of expertise and training in the detection of concealed firearm carrier through mock CCTV-footage. CCTV operators did not differ from lay people in their ability to identify a carrier of a concealed firearm. Furthermore, these two groups of participants did not differ in their performance on the test of non-verbal sensitivity. Additionally, the test of non-verbal sensitivity shows not enough relationship with performance on affect detection and gun carrier detection tasks to be able to conclude that the decoding ability of observers measured by the PONS test can predict this performance. In future research another way to measure this relationship, perhaps in the form of a new screening tool more suitable for the current purpose, should be developed.

In accordance with previous findings about the possible superiority of the frontal view in recognition of human actions and affect, the research showed that viewing the footage of surveillance targets from a street-level perspective compared to typical CCTV-like perspective may increase awareness of affective state of surveillance targets and enhance the accuracy with which an observer can judge the emotional state of another. The present research has also identified several cues associated with concealed firearm carrying that can be used in further research. Those cues were found to be associated with the performance of human observers on two tasks: affect detection task and gun carrier detection task. Applying the questionnaire and the eye movement registration provided information about the way the observers were using such cues as gait, posture, arm movements, the face, upper body and lower body of surveillance targets.

Taken together, the results of the present research showed that the recognition of concealed firearm carrying is highly ambiguous in terms of visual detection of it, and that those engaged in a surveillance task might use visual indicators of affective state of surveillance targets to make a decision whether or not the individuals are carrying a concealed firearm. Although it cannot be inferred that the use of solely visual indicators of affective state would support
reliable detection of concealed firearm carrying, it should be especially useful to consider the affective behavioural cues when also other suspicious visual cues (e.g., gun shaped bulge in the clothing) are visible to an observer. In this respect it would be interesting to further investigate how the level of certainty in the decision making process concerning the detection of concealed firearm carrying would vary depending on either present or absent information about affective behavioural cues. Moreover, further investigation is needed into the possibility to generate a set of strategies that would employ affective behavioural cues and would lead to more reliable detection of concealed firearm carrying.
REFERENCES


Donald, C.H.M., 2006. Learn to see the body talk. CCTV image. CCTV User Group, p. 28.


Harvey, L.O. Jr., 2001. ‘*RSCORE 5.3.2*”. Department of Psychology, University of Colorado, Boulder, Colorado.


Keval, H. and Sasse, M.A., 2008 -2. Can we ID from CCTV? Image quality in
digital CCTV and face identification performance. SPIE Mobile
Multimedia/Image Processing, Security, and Applications, Agaian, S.S., Jassim,

Keval, H.U. and Sasse, M.A., 2010. 'Not the usual suspects': A study of factors

Klinesmith, J., Kasser, T., McAndrew, F., 2006. Guns, testosterone, and
aggression: an experimental test of a meditational hypothesis. Psychological
Science, 179, pp. 568-571.

Kohler, E., Keysers, C., Umilta’, M.A., Fogassi, L., Gallese, V., Rizzolatti, G.,
2002. Hearing sounds, understanding actions: Action representation in mirror

Kozlowski, L. T. and Cutting, J. E., 1977. Recognizing the sex of a walker from

Crime prevention studies, 1, pp. 55-68.

Note 2, September, (E&W).

adjective check list – revised. Educational and Industrial Testing Service, San
Diego, California.

Luff, P., Heath, C., and Jirotka, M ., 2000. Surveying the Scene: Technologies
for Everyday Awareness and Monitoring in Control Rooms. Interacting with


APPENDIX 1

AGGRESSION QUESTIONNAIRE (BUSS, A.H. and PERRY, M., 1992)

Instructions: Using the 5-point scale shown below, indicate how uncharacteristic or characteristic each of the following statements is in describing you.

Place your rating in the box to the right of the statement.

1 = extremely uncharacteristic of me
2 = somewhat uncharacteristic of me
3 = neither uncharacteristic nor characteristic of me
4 = somewhat characteristic of me
5 = extremely characteristic of me

<table>
<thead>
<tr>
<th>Rating</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Some of my friends think I am a hothead</td>
</tr>
<tr>
<td>2</td>
<td>If I have to resort to violence to protect my rights, I will.</td>
</tr>
<tr>
<td>3</td>
<td>When people are especially nice to me, I wonder what they want.</td>
</tr>
<tr>
<td>4</td>
<td>I tell my friends openly when I disagree with them.</td>
</tr>
<tr>
<td>5</td>
<td>I have become so mad that I have broken things.</td>
</tr>
<tr>
<td>6</td>
<td>I can’t help getting into arguments when people disagree with me.</td>
</tr>
<tr>
<td>7</td>
<td>I wonder why sometimes I feel so bitter about things.</td>
</tr>
<tr>
<td>8</td>
<td>Once in a while, I can’t control the urge to strike another person.</td>
</tr>
<tr>
<td>9</td>
<td>I am an even-tempered person.</td>
</tr>
<tr>
<td>10</td>
<td>I am suspicious of overly friendly strangers.</td>
</tr>
<tr>
<td>11</td>
<td>I have threatened people I know.</td>
</tr>
<tr>
<td>12</td>
<td>I flare up quickly but get over it quickly.</td>
</tr>
<tr>
<td>13</td>
<td>Given enough provocation, I may hit another person.</td>
</tr>
<tr>
<td>14</td>
<td>When people annoy me, I may tell them what I think of them.</td>
</tr>
<tr>
<td>15</td>
<td>I am sometimes eaten up with jealousy.</td>
</tr>
<tr>
<td>16</td>
<td>I can think of no good reason for ever hitting a person.</td>
</tr>
<tr>
<td>17</td>
<td>At times I feel I have gotten a raw deal out of life.</td>
</tr>
<tr>
<td>18</td>
<td>I have trouble controlling my temper.</td>
</tr>
<tr>
<td>19</td>
<td>When frustrated, I let my irritation show.</td>
</tr>
<tr>
<td>20</td>
<td>I sometimes feel that people are laughing at me behind my back.</td>
</tr>
<tr>
<td>21</td>
<td>I often find myself disagreeing with people.</td>
</tr>
<tr>
<td>22</td>
<td>If somebody hits me, I hit back.</td>
</tr>
<tr>
<td>23</td>
<td>I sometimes feel like a powder keg ready to explode.</td>
</tr>
<tr>
<td>24</td>
<td>Other people always seem to get the breaks.</td>
</tr>
<tr>
<td>25</td>
<td>There are people who pushed me so far that we came to blows.</td>
</tr>
<tr>
<td>26</td>
<td>I know that “friends” talk about me behind my back.</td>
</tr>
<tr>
<td>27</td>
<td>My friends say that I’m somewhat argumentative.</td>
</tr>
<tr>
<td>28</td>
<td>Sometimes I fly off the handle for no good reason.</td>
</tr>
<tr>
<td>29</td>
<td>I get into fights a little more than the average person.</td>
</tr>
</tbody>
</table>
## APPENDIX 2

### MAACL-R STATE FORM (LUBIN, B. and ZUCKERMAN, M., 1999)

Participant number: On this sheet you will find words which describe different moods and feelings. Mark an X beside the words that describe how YOU FEEL AT THIS MOMENT. Some words may sound alike, but we want you to CHECK ALL THE WORDS that DESCRIBE your feelings. Please read EVERY word and consider whether it applies to you or not. Work rapidly.

<table>
<thead>
<tr>
<th>active</th>
<th>fit</th>
<th>peaceful</th>
</tr>
</thead>
<tbody>
<tr>
<td>adventurous</td>
<td>forlorn</td>
<td>pleased</td>
</tr>
<tr>
<td>affectionate</td>
<td>frank</td>
<td>pleasant</td>
</tr>
<tr>
<td>afraid</td>
<td>free</td>
<td>polite</td>
</tr>
<tr>
<td>agitated</td>
<td>friendly</td>
<td>powerful</td>
</tr>
<tr>
<td>agreeable</td>
<td>frightened</td>
<td>quiet</td>
</tr>
<tr>
<td>aggressive</td>
<td>furious</td>
<td>reckless</td>
</tr>
<tr>
<td>alive</td>
<td>lively</td>
<td>rejected</td>
</tr>
<tr>
<td>alone</td>
<td>gentle</td>
<td>rough</td>
</tr>
<tr>
<td>amiable</td>
<td>glad</td>
<td>sad</td>
</tr>
<tr>
<td>amused</td>
<td>gloomy</td>
<td>safe</td>
</tr>
<tr>
<td>angry</td>
<td>good</td>
<td>satisfied</td>
</tr>
<tr>
<td>annoyed</td>
<td>good-natured</td>
<td>secure</td>
</tr>
<tr>
<td>awful</td>
<td>grim</td>
<td>shabby</td>
</tr>
<tr>
<td>bashful</td>
<td>happy</td>
<td>shaky</td>
</tr>
<tr>
<td>bitter</td>
<td>healthy</td>
<td>soothed</td>
</tr>
<tr>
<td>blue</td>
<td>hopeless</td>
<td>steady</td>
</tr>
<tr>
<td>bored</td>
<td>hostile</td>
<td>stubborn</td>
</tr>
<tr>
<td>calm</td>
<td>impatient</td>
<td>stormy</td>
</tr>
<tr>
<td>cautious</td>
<td>incensed</td>
<td>strong</td>
</tr>
<tr>
<td>cheerful</td>
<td>indignant</td>
<td>suffering</td>
</tr>
<tr>
<td>clean</td>
<td>inspired</td>
<td>sullen</td>
</tr>
<tr>
<td>complaining</td>
<td>interested</td>
<td>sunk</td>
</tr>
<tr>
<td>contented</td>
<td>irritated</td>
<td>sympathetic</td>
</tr>
<tr>
<td>contrary</td>
<td>jealous</td>
<td>tame</td>
</tr>
<tr>
<td>cool</td>
<td>joyful</td>
<td>tender</td>
</tr>
<tr>
<td>cooperative</td>
<td>kindly</td>
<td>tense</td>
</tr>
<tr>
<td>critical</td>
<td>lonely</td>
<td>terrible</td>
</tr>
<tr>
<td>cross</td>
<td>lost</td>
<td>terrified</td>
</tr>
<tr>
<td>cruel</td>
<td>loving</td>
<td>thoughtful</td>
</tr>
<tr>
<td>daring</td>
<td>low</td>
<td>timid</td>
</tr>
<tr>
<td>desperate</td>
<td>lucky</td>
<td>tormented</td>
</tr>
<tr>
<td>destroyed</td>
<td>mad</td>
<td>understanding</td>
</tr>
<tr>
<td>devoted</td>
<td>mean</td>
<td>unhappy</td>
</tr>
<tr>
<td>disagreeable</td>
<td>meek</td>
<td>unsociable</td>
</tr>
<tr>
<td>discontented</td>
<td>merry</td>
<td>upset</td>
</tr>
<tr>
<td>discouraged</td>
<td>mild</td>
<td>vexed</td>
</tr>
<tr>
<td>disgusted</td>
<td>miserable</td>
<td>warm</td>
</tr>
<tr>
<td>displeased</td>
<td>nervous</td>
<td>whole</td>
</tr>
<tr>
<td>energetic</td>
<td>obilging</td>
<td>wild</td>
</tr>
<tr>
<td>enraged</td>
<td>offended</td>
<td>wilful</td>
</tr>
<tr>
<td>enthusiastic</td>
<td>outraged</td>
<td>wilted</td>
</tr>
<tr>
<td>fearful</td>
<td>panicky</td>
<td>worrying</td>
</tr>
<tr>
<td>fine</td>
<td>patient</td>
<td>young</td>
</tr>
</tbody>
</table>
APPENDIX 3

**MAACL-R TRAIT FORM** *(LUBIN, B. and ZUCKERMAN, M., 1999)*

Participant number: On this sheet you will find words which describe different moods and feelings. Mark an X beside the words that describe how YOU GENERALLY FEEL, that is how you have GENERALLY FELT OVER THE PAST MONTH. Some words may sound alike, but we want you to CHECK ALL THE WORDS that DESCRIBE your feelings. Please read EVERY word and consider whether it applies to you or not. Work rapidly.

<table>
<thead>
<tr>
<th>active</th>
<th>fit</th>
<th>peaceful</th>
</tr>
</thead>
<tbody>
<tr>
<td>adventurous</td>
<td>forlorn</td>
<td>pleased</td>
</tr>
<tr>
<td>affectionate</td>
<td>frank</td>
<td>pleasant</td>
</tr>
<tr>
<td>afraid</td>
<td>free</td>
<td>polite</td>
</tr>
<tr>
<td>agitated</td>
<td>friendly</td>
<td>powerful</td>
</tr>
<tr>
<td>agreeable</td>
<td>frightened</td>
<td>quiet</td>
</tr>
<tr>
<td>aggressive</td>
<td>furious</td>
<td>reckless</td>
</tr>
<tr>
<td>alive</td>
<td>lively</td>
<td>rejected</td>
</tr>
<tr>
<td>alone</td>
<td>gentle</td>
<td>rough</td>
</tr>
<tr>
<td>amiable</td>
<td>glad</td>
<td>sad</td>
</tr>
<tr>
<td>amused</td>
<td>gloomy</td>
<td>safe</td>
</tr>
<tr>
<td>angry</td>
<td>good</td>
<td>satisfied</td>
</tr>
<tr>
<td>annoyed</td>
<td>good-natured</td>
<td>secure</td>
</tr>
<tr>
<td>awful</td>
<td>grim</td>
<td>shaky</td>
</tr>
<tr>
<td>bashful</td>
<td>happy</td>
<td>shy</td>
</tr>
<tr>
<td>bitter</td>
<td>healthy</td>
<td>soothed</td>
</tr>
<tr>
<td>blue</td>
<td>hopeless</td>
<td>steady</td>
</tr>
<tr>
<td>bored</td>
<td>hostile</td>
<td>stubborn</td>
</tr>
<tr>
<td>calm</td>
<td>impatient</td>
<td>stormy</td>
</tr>
<tr>
<td>cautious</td>
<td>incensed</td>
<td>strong</td>
</tr>
<tr>
<td>cheerful</td>
<td>indignant</td>
<td>suffering</td>
</tr>
<tr>
<td>clean</td>
<td>inspired</td>
<td>sullen</td>
</tr>
<tr>
<td>complaining</td>
<td>interested</td>
<td>sunk</td>
</tr>
<tr>
<td>contented</td>
<td>irritated</td>
<td>sympathetic</td>
</tr>
<tr>
<td>contrary</td>
<td>jealous</td>
<td>tame</td>
</tr>
<tr>
<td>cool</td>
<td>joyful</td>
<td>tender</td>
</tr>
<tr>
<td>cooperative</td>
<td>kindly</td>
<td>tense</td>
</tr>
<tr>
<td>critical</td>
<td>lonely</td>
<td>terrible</td>
</tr>
<tr>
<td>cross</td>
<td>lost</td>
<td>terrified</td>
</tr>
<tr>
<td>cruel</td>
<td>loving</td>
<td>thoughtful</td>
</tr>
<tr>
<td>daring</td>
<td>low</td>
<td>timid</td>
</tr>
<tr>
<td>desperate</td>
<td>lucky</td>
<td>tormented</td>
</tr>
<tr>
<td>destroyed</td>
<td>mad</td>
<td>understanding</td>
</tr>
<tr>
<td>devoted</td>
<td>mean</td>
<td>unhappy</td>
</tr>
<tr>
<td>disagreeable</td>
<td>meek</td>
<td>unsociable</td>
</tr>
<tr>
<td>discontented</td>
<td>merry</td>
<td>upset</td>
</tr>
<tr>
<td>discouraged</td>
<td>mild</td>
<td>vexed</td>
</tr>
<tr>
<td>disgusted</td>
<td>miserable</td>
<td>warm</td>
</tr>
<tr>
<td>displeased</td>
<td>nervous</td>
<td>whole</td>
</tr>
<tr>
<td>energetic</td>
<td>offended</td>
<td>wild</td>
</tr>
<tr>
<td>enraged</td>
<td>outraged</td>
<td>wilted</td>
</tr>
<tr>
<td>enthusiastic</td>
<td>panicked</td>
<td>worrying</td>
</tr>
<tr>
<td>fearful</td>
<td>patient</td>
<td>young</td>
</tr>
</tbody>
</table>
APPENDIX 4

CUE-DETECTION QUESTIONNAIRE

Question 1
Please describe the overall movement of actors using the following characteristics of actors’ movement pattern (tick the appropriate answers):

Actor
uses short strides uses medium strides length uses long strides (length of stride)
does not swing arms does swing arms a little swings arms a lot (degree of arms’ swinging)
was light-footed was neither light - nor heavy-footed was very heavy-footed (degree of heavy footedness)

The movements were
smooth neither smooth, nor jerky jerky (degree of regularity)
stiff neither stiff, nor loose loose (degree of reticence)
soft neither soft, nor hard hard (degree of suppleness)
slow neither slow, nor fast fast (degree of speed)
expanded neither expanded, nor contracted contracted (degree of tension)

Question 2
In the video clips you have just seen, how did you gain an impression of the moods of the people in the clips? Please rate each of the following statements

I looked at their gait
not at all rarely sometimes often all the time

I looked at their posture
not at all rarely sometimes often all the time

I looked at their facial expressions
not at all rarely sometimes often all the time

I looked at the position of their arms
not at all rarely sometimes often all the time
APPENDIX 5

PARTICIPANT INFORMATION SHEET

This is an invitation to take part in a study to see how well you can detect mood of
people who were filmed for the experimental purposes.

The purpose of the study

The purpose of this study is to find out how well a naïve observer can identify
emotional state of others observed in staged CCTV footage.

The information gathered from this experiment will be used for two purposes:
to help train CCTV operators and other security professionals to detect firearms;
to further scientific knowledge by publishing anonymous data in academic journals and
at conferences.

Taking part
Taking part will involve the following steps.
Answer a few questions about your personal information;
Complete a computer-based experiment designed to work out how well you can detect
emotional state of someone from staged CCTV-footage

This whole process should take no longer than an hour. You can take a break
whenever you like.

You will be paid £5 inconvenience allowance for your participation in this experiment.

You will be free to quit the experiment at any time. You do not have to explain your
reasons for quitting to the experimenter. The experimenter will be on hand to discuss
any issues that you may have with regard to the experiment.

Your data

The only information we wish keep from the study are the answers you give in the
short interview and the responses you give during the computer-based experiments.
This information will not be associated with your name. It will be stored securely.
Where your data is stored in hardcopy, it will be locked in a filing cabinet to which
access is restricted to the research team. Where your data is stored electronically, it
will be saved on a university owned PC, in password protected files, to which access is
restricted to the research team. Your data will be stored in accordance with the Data
Protection Act. You can request that your data be destroyed at any time.
APPENDIX 6

INFORMED CONSENT FORM
(to be completed after Participant Information Sheet has been read)

The purpose and details of this study have been explained to me.
I understand that this study is designed to further scientific knowledge and that all procedures have been approved by the Loughborough University Ethical Advisory Committee.
I have read and understood the information sheet and this consent form.
I have had an opportunity to ask questions about my participation.
I understand that I am under no obligation to take part in the study.
I understand that I have the right to withdraw from this study at any stage for any reason, and that I will not be required to explain my reasons for withdrawing.
I understand that all the information I provide will be treated in strict confidence.
I agree to participate in this study.

Your name

Your signature

Signature of investigator

Date
APPENDIX 7

PARTICIPANTS' INSTRUCTIONS FOR EXPERIMENT 'MEDUSA: Eye-movements associated with recognition of mood in people observed via CCTV'

Instructions:
It is known that carrying a firearm might change the mood and the non-verbal behaviour (i.e. body movement, posture, gait, facial expressions etc.) of the carrier. It is also known that carrying a firearm can cause the negative feelings by the carrier, as for example, anxiety, hostility or depression. As result of it the posture or the way of walking of a gun carrier can appear to observers as anxious, hostile or depressed.

In the following experiment you will see 22 short video clips of various people walking along a corridor. We are interested in whether or not you think they were carrying a concealed gun.

Each clip will show just one person. The person will be carrying either a gun or a bottle of soft drink, hidden somewhere on the person (e.g., in a pocket, under a jacket, up a sleeve, in a hood, in a sock, in the waistband of the trousers or jeans).

If they are carrying a gun it will be a semi-automatic pistol.

If they are carrying a bottle of drink it will be a one-litre bottle of soft drink.

After each clip has finished playing you will see six options. Please indicate whether or not you think the person in the clip that you have just been watching was carrying a gun, and indicate how confident you are about your answer by selecting one of these options:
Definitely NO   Probably NO   At a guess, NO   At a guess, YES   Probably YES   Definitely YES

If you think that they were carrying a gun please select one of the "yes" answers. However, if you think that they were carrying a bottle, please select one of the "no" answers. When you have selected an answer, a 'Submit' button will appear. Before you click on 'Submit' button you are able to change your answer, but you are encouraged to go with your first answer or gut feeling. After you pressed the 'Submit' button, the next clip will be shown immediately, so please be ready for it.
DEBRIEFING SHEET

Thank you for taking part in our study. This information sheet is just to recap the purposes of the study and to tell you how the research might progress.

The aim of this study:
help train CCTV operators and other security professionals to detect firearms;
to acquire scientific knowledge in the security field by publishing anonymised data in academic journals and conferences.

In this study we have used staged footage of people carrying guns, as real-life footage of gun crime is rare and difficult to get hold of. This experiment is a first attempt to understand problems related to spotting a gun. We are currently collecting as much real-life footage of gun crime as possible from CCTV control rooms and police forces. In future studies we hope to use this real-life footage of gun crime in a similar experiment in order to achieve a greater level of real-world validity.

At this point we would like to remind you that you can ask for your data to be destroyed at any time. You do not have to explain your reasons for requesting that your data be destroyed.

Please feel free to contact the research team about any issues relating to the study: