Detecting Deception Through Non-Verbal Behaviour

by

Tessa Catherine Mapala

A thesis submitted in partial fulfilment of the requirements of Lancaster University

for the degree of Doctor of Philosophy

Lancaster University

December 2020

Declaration

I declare that this thesis is my own work and has not been submitted in the same form for the award of a higher degree elsewhere.

Abstract

The security protocols used in airport security checkpoints primarily aim to detect prohibited items, as well as the detection of malicious intent and associated deception to thwart any threats. However, some of the security protocols that are used are not substantiated by scientifically validated cues of deception. Instead, some protocols, such as the Screening of Passengers by Observation Techniques (SPOT) program, have been developed based on anecdotal evidence and invalid cues of deception. As such, the use of these protocols has received a lot of criticism in recent years from government agencies, civil rights organisations and academia. These security protocols rely on security personnel's ability to infer intent from non-verbal behaviour, yet the literature suggests that the relationship between non-verbal cues and deception is unreliable and that people are poor at detecting deception. To improve upon our understanding of the validity of these protocols, this thesis used virtual reality to replicate a security checkpoint to explore whether there were valid cues of deception, specifically in an airport context. People's ability to identify whether others were behaving deceptively was assessed, as well as the factors that may be informing decision-making.

Chapter Four of this thesis found that the non-verbal cues of interest, which were segment displacement, centre of mass displacement, cadence, step length and speed were not significantly different between honest and deceptive people. A verbal measure, response latency, was found to only distinguish between honest people and those who were deceptive about a future intention, but not those who were deceptive about having a prohibited item.

In light of the use of non-verbal measures in practice despite the lack of scientific support, Chapters Five to Seven aimed to gain a greater insight into people's deception detection capabilities. The findings from Chapters Five to Seven reflected that the ability to detect deception from non-verbal behaviour was no better than guessing. Specifically,

Chapter Five found that the accuracy of detecting deception was no different from chance levels. Six themes emerged as the factors that were used to inform decision-making. The themes were physical appearance, disposition, walking behaviour, body positioning, looking behaviour and upper limb movement, though a qualitative analysis revealed that there were subjective interpretations of how the themes mapped onto deception. Chapter Six introduced two techniques of information reduction to assess whether accuracy could be improved above chance levels by lessening the impact of biasing factors. Neither technique resulted in accuracy above chance levels. In Chapter Seven, eye tracking was utilised to assess the gaze patterns associated with the detection of deception. People looked at the legs more than other areas of the body prior to decision-making, though only looking at the left arm and hand were linked with accuracy. Detection accuracy was poor overall, though looking at the left arm was linked with reduced accuracy, whilst looking at the left hand was linked with increased accuracy.

Overall, this thesis showed that the non-verbal cues that were assessed could not distinguish between honest and deceptive people. In the absence of valid cues, observers were not able to identify deception at a rate above chance even with the reduction of potentially biasing factors. The results of this thesis reinforce the idea that incorporating non-verbal measures into threat/deception detection protocols may not be warranted because of the dubious nature of their reliability and validity, as well as the poor deception identification capabilities when relying on non-verbal behaviour.

Table of Contents

Declaration ii

Abstract iii

Table of Contents v

List of Tables ix

List of Figures x

Part I: Chapter One- Thesis outline 1

Chapter Two- Literature review 5

2.1 General introduction 5

2.1.1 Non-verbal cues in practice 7

2.2 Displaying non-verbal cues of deception 10

2.2.1 The multi-factor theory’s view of non-verbal deceptive behaviour 10

2.2.2 The self-presentation theory’s view of non-verbal deceptive behaviour 17

2.3 Identifying non-verbal cues of deception 19

2.3.1 Theoretical approaches to identifying deception 21

2.3.2 Poor accuracy detecting deception 25

2.3.3 The research-practice gap 29

2.4 Making advances in non-verbal deception detection with technology 31

2.4.1 Contextually salient research methods 32

2.4.2 Manual versus automated methods of measuring non-verbal behaviour 34

2.5 Current thesis 38

Chapter Three- Methodology 40

3.1 Motion capture 40

3.1.1 Recording behaviour using the Xsens system 41

3.1.2 Processing the non-verbal data 42

3.1.3 Analysing the non-verbal data 42

3.2 Virtual reality 46

3.2.1 Virtual reality methodology in this thesis 49

3.3 Point-light displays 53

3.4 Eye-tracking 54

Part II: Chapter Four – Investigating Deception Cues 56

4.1 Introduction 56

4.2 Method 64

4.2.1 Participants 64

4.2.2 Sampling 66

4.2.3 Materials 66
4.2.4 Measures 67
4.2.5 Design 69
4.2.6 Procedure 70
4.2.7 Analysis 72
4.3 Results 73
4.3.1 Non-verbal results 73
4.3.2 Response latency results 77
4.4 Discussion 79
4.4.1 Non-verbal behaviour discussion 79
4.4.2 Response latency discussion 80
4.4.3 Limitations 81
4.4.4 Practical application 84
4.4.5 Summary 85
Part III: Chapter Five – Deception Identification Accuracy and Decision-Making 86
5.1 Introduction 86
5.2 Methods 89
5.2.1 Participants 89
5.2.2 Materials 90
5.2.3 Measures 91
5.2.4 Design 91
5.2.5 Procedure 91
5.2.6 Analysis 92
5.3 Results 93
5.3.1 Accuracy 93
5.3.2 Thematic analysis 93
5.4 Discussion 98
5.4.1 Accuracy commentary 98
5.4.2 Thematic analysis commentary 99
5.4.3 Limitations 101
5.4.4 Summary 102
Chapter Six - Information Reduction Techniques and Decision Accuracy 103
6.1 Introduction 103
6.2 Methods 110
6.2.1 Participants 110
6.2.2 Materials 111
6.2.3 Measures 112

6.2.4 Design 112
6.2.5 Procedure 113
6.2.6 Analysis 116
6.3 Results 116
6.4 Discussion 117
6.4.1 Summary 119
Part IV: Chapter Seven – Gaze Behaviour Before Deception Identification 120
7.1 Introduction 120
7.2 Eye-tracking and deception research 120
7.3 Methods 122
7.3.1 Participants 122
7.3.2 Materials 123
7.3.3 Measures 124
7.3.4 Design 125
7.3.5 Procedure 125
7.3.6 Analysis 126
7.4 Results 126
7.5 Discussion 127
7.5.1 Limitations 129
7.5.2 Summary 129
Part V: Chapter Eight- General Discussion 130
8.1 Overview 130
8.2 The use of technology in this research 131
8.2.1 Virtual reality 131
8.2.2 Measuring non-verbal behaviour 134
8.3 Empirical summaries and theoretical implications 135
8.3.1 Deception cues 135
8.3.2 Accuracy identifying deception and the factors that influence decision-making 139
8.3.3 The effect of reducing subjective influence on decision accuracy 141
8.3.4 Eye gaze behaviour 143
8.3.5 Implications summary 144
8.4 Practical implications 145
8.5 Ethical implications 147
8.6 Overall limitations 147
8.7 Concluding remarks 149
References 150

Appendix A- Avatar samples 180
Appendix B- Immersion Questionnaire 183
Appendix C- Questions asked in Chapter Four 185

List of Tables

Table 4.1 Segment displacement MANOVA comparing the honest condition to the three deceptive conditions 73

Table 4.2 Segment displacement MANOVA comparing the future action condition to the package and weapon conditions 74

Table 4.3 Segment displacement MANOVA comparing the package and weapon condition 75

Table 4.4 Means and standard deviations of the segment displacement between conditions 75

Table 4.5 Means and standard deviations of movement per minute, cadence, step length, speed and centre of mass displacement 76

Table 4.6 Linear mixed model with condition (honest condition as the intercept) and response type (honest response as the reference level) as predictor variables 77

Table 4.7 Random effects for the linear mixed model in Table 4.6 77

Table 4.8 Descriptive summary of the response latencies with the response type 78

Table 5.1 Binomial logistic regression results with avatar deceptive condition as the predictor variable (honest as the reference level), the time spent viewing the avatars as a covariate and accuracy as the dependent variable. 93

Table 6.1 Results of the logistic regression, with accuracy as the dependent variable 116

List of Figures

Figure 2.1 A Brunswik's lens model of deception detection. 20

Figure 3.1 Xsens model depicting the sensors placements and segments for analysis 40

Figure 3.2 Virtual environment from the point of view of a participant in the role of a passenger as in Chapter Four. 51

Figure 3.3 Virtual environment from the point of view of a participant in the role of security personnel, as in Chapters Five, Six and Seven. 51

Figure 3.4 Comparison of Xsens representation and a point-light display. 53

Figure 3.5 Diagram of the eye-tracker setup. 55

Figure 4.1 Five key stages of the experimental procedure whilst in virtual reality, during which non-verbal behaviour and response latency were measured 72

Figure 4.2 Violin plot of the response latencies for the honest, future action, package and weapon condition 78

Figure 6.1 Study flow with both the reduced superficial information and imposed time-constraints condition 115

Figure 7.1 An avatar with six areas of interest depicted. 124

Part I: Chapter One- Thesis outline

Aspects of non-verbal behaviour are commonly used to judge the intent or inner states of other people. Notably, within some airport security protocols, security personnel are tasked with assessing people's non-verbal behaviour for traits that signify that they pose a threat or may be deceptive about the nature of their intentions. These assessments of non-verbal behaviour are used to justify approaching people for further questioning (U.S. Government Accountability Office, 2010; Winter & Currier, 2015). Assessments of non-verbal behaviour as ascribed by some security protocols (and the stop and search police procedure) have been plagued by criticisms of their validity and the resultant occurrence of ethnic, racial, gender and religious discrimination when determining who may pose a threat (ACLU, 2017; Bradford & Tiratelli, 2019; Homeland Security Digital Library, 2002; Meyer, 2010).

Importantly, decades of research have failed to find reliable and valid non-verbal behavioural markers which distinguish can identify deception. Coupled with the findings that people are poor at detecting deception in others, especially when using non-verbal methods alone (Vrij et al., 2019), it would appear that non-verbal measures may not be appropriate for application in airport settings at present.

The pseudoscientific (invalid) nature of some security protocols is an important problem to tackle because decision-making in this setting that is not grounded in scientific findings could lead to inconsistent judgements that are informed by prejudicial stereotypes. As a result, honest people may be incorrectly determined to be deceptive and people who are deceptive but do not conform to a stereotype could go unnoticed (Denault et al., 2020; Lilienfeld & Landfield, 2008; Murai et al., 2018), both of which could have a detrimental impact on an individuals' rights and security, respectively.

One potential reason why pseudoscientific methods are being incorporated into security protocols (Denault & Jupe, 2018) is that some practitioners may have concerns that laboratory studies on deception do not closely reflect security environments, thus potentially affecting the behaviours exhibited, as well as the utility and desire to appropriately transfer the findings to the real world (Buckley, 2012; Denault & Jupe, 2018). Within the deception field, few studies are contextually relevant to an airport environment which may limit the ability to determine the transferability of research on non-verbal deception to a security checkpoint context. A focus of deception research needs to be on designing studies with the aim of bridging the gap between the laboratory findings and the real world to inch toward gaining a better understanding of non-verbal behaviour in settings where deception can have important consequences. By incorporating environments where non-verbal assessments are primarily used, such as airports (Vrij et al., 2019) into research, we may contribute toward lessening the reliance on pseudoscientific measures and the development of more effective and scientifically valid protocols. This thesis aims to address this problem by using virtual reality to simulate a security checkpoint to assess the validity of some non-verbal measures in this context. Virtual reality is used to induce contextual priming (Cerda et al., 2021), whereby the familiarity of the visual information may activate behavioural responses akin to being at a security checkpoint.

This thesis consists of five parts and is structured to approach deception detection from two sides: the person attempting the deception (Part II) and the person who is tasked with detecting deception (Part III and IV).

<u>Part I</u>

Within Part I, there are three chapters. This introductory chapter outlines the aims and structure of this thesis. Chapter Two presents a literature review on some of the theoretical views within the field of non-verbal measures of deception detection. The applied use of deception detection at security checkpoints is discussed, as well as the critiques that some deception detection protocols have garnered over the past few decades. The relationship between academia and the real world is also briefly explored, as well as how technology can be used within non-verbal deception research. Chapter Three details the four main types of technology that are used in the empirical chapters of this thesis. These technologies include the use of motion capture (Chapters Four to Seven), virtual reality (Chapter Four to Chapter Seven), point-light displays (Chapter Six) and eye-tracking (Chapter Seven).

Part II

Objective 1: Investigate whether there are any scientifically valid cues for detecting deception.

Chapter Four aims to answer the first main objective by assessing whether there are valid non-verbal deception markers in the context of a virtual security checkpoint. The motion capture data gained from participants in Chapter Four is later used in Chapters Five, Six and Seven (details of this are expounded on in Chapter Three). A conceptual replication of a prior verbal response latency study (Mapala et al., 2017) is also included in Chapter Four.

Part III

Objective 2. Investigate people's accuracy at detecting deception.

Objective 3. Improve accuracy by testing two strategies of information reduction: reduced superficial information and imposed time-constraints

There are two chapters, Five and Six, within this section, both of which aim to assess the second objective: people's accuracy in detecting deception. Chapter Six also assesses the third objective and builds on Chapter Five by employing two different techniques of information reduction to see whether accuracy can be improved above chance levels.

Part IV

Objective 4. Investigate what people look at prior to deciding whether deceit is present.

Chapter Seven investigates the eye gaze behaviour of people when they are assessing whether others are deceptive.

Part V

Chapter Eight provides an overall summary of the thesis and the validity of non-verbal measures to identify deception. This chapter also discusses the use of virtual reality and motion capture in this thesis, the potential for future research, the theoretical, practical and ethical implications for deception detection and the limitations of this thesis.

Chapter Two- Literature Review

2.1 General introduction

This literature review will show that given the unpredictable nature of the exhibition of non-verbal cues and poor ability to detect deception, the incorporation of non-verbal measures in security procedures is premature, unsubstantiated by scientific evidence and may lead to judgements based on factors irrelevant to deception. The context of previous studies on deception will also be discussed in relation to the potential difficulty of generalising their findings to our understanding of deception specifically in an airport context. There are three overarching sections of this literature review which will discuss: (i) how, when and why non-verbal behaviour may change when engaging in deception (section 2.2); (ii) the ability to detect deception and the factors that are a detriment to accurately identifying deception (section 2.3); (iii) how technological interventions may aid our understanding of non-verbal deception (section 2.4).

Numerous definitions of deception exist (Carson, 2001). This thesis defines deception as instances of a communicator deliberately conveying false information by fabrication, concealment, or manipulation, with the intention that others will believe it to be true (Gorski, 1998; Mahon, 2016). This interpretation is similar to the Cambridge dictionary definition of deception as “a statement or action that hides the truth” (Cambridge University Press, 2020). In the broader scheme of things, deception is “a process, [and] lying refers to a specific strategy that can be used within this process” (Ströfer et al., 2015, p. 1). However, these two terms (deception and lying) are used interchangeably within the literature and within this thesis.

Attempts to establish scientifically valid methods of detecting deception came into prominence in the late 19th century with phrenology, which though now discredited (Gross, 2001), spurred the idea that deception could and should be investigated scientifically (Vicianova & Jozef, 2015). This movement was followed by graphology (Ford, 2006), the polygraph test (Horvath & Reid, 1971) and the present focus on verbal and non-verbal cues (Vrij et al., 2010; Zuckerman et al., 1981). Verbal cues refer to aspects of verbal communication that may signal that someone is being deceptive, such as less detailed statements (Granhag & Strömwall, 2002), inconsistencies with evidence (Hartwig et al., 2011) and increased response latencies (Suchotzki et al., 2013). Non-verbal cues refer to differences in body movements that may signal that someone is being deceptive, for example, reduced leg movements (Sporer & Schwandt, 2007), fewer illustrative gestures (Ekman et al., 1976) and increased postural rigidity (Twyman et al., 2014). For the purpose of this thesis, non-verbal behaviour is limited to physical movements, although other research has sometimes included aspects of speech, such as pitch (Vrij & Winkel, 1991).

Non-verbal deception detection will be the primary focus of this thesis, except for a response latency measure in Chapter Four as a conceptual replication of a previous study. In this literature review chapter, verbal behaviour will only be discussed directly concerning non-verbal measures of deception, such as the impact of the cognitive interview on non-verbal behaviour in section 2.2.1.1, but not as a standalone measure of deception. There are a few reasons for the focus on non-verbal measures in this thesis. When judging the intentions of others, people tend to default to assessing non-verbal behaviour to inform their decisions, for reasons such as believing that it is difficult for others to prevent conveying their true intentions with their non-verbal behaviour (Eubanks et al., 2015). Unlike verbal communication, where one can simply cease speaking, non-verbal behaviour can be more difficult to mute because your body is always 'talking' and conveying some form of information (Eubanks et al., 2015), which may explain why people rely on it. For example, the kinematic specification of dynamics principle (Runeson & Frykholm, 1983) proposes that we each have a 'kinematic fingerprint'. Similar to a regular fingerprint, this principle posits that gait is unique to each person and importantly, that deception can be inferred from kinematic behaviour alone. This idea was informed by a series of point-light studies, which will be discussed further in Chapter Six. Sometimes, defaulting to judging people based on their non-verbal behaviour may be for practical reasons. In some instances, such as when judging people from afar in an airport environment, non-verbal behaviour may be the only information source available.

2.1.1 Non-verbal cues in practice

After 9/11, global security concerns dramatically increased, particularly in the aviation industry. In the USA, the Transport Security Administration (TSA) was formed in November of that year, operating at airports in the United States (Blalock et al., 2007). In the UK, the 7/7 bombings displayed London's vulnerability to attack on the underground and bus networks and were followed by the introduction of the Terrorism Act (Parker, 2007). More recently, attacks on airports in Brussels and Istanbul in 2016, and the Manchester arena in 2017 have highlighted the threat that is potentially posed to such venues and has shone a light on the need for more research to bolster attempts at thwarting such tragedies.

Passive observational assessments of non-verbal behaviour can be the primary driver prompting security personnel to approach people, both in environments such as airports with the United States' TSA Screening of Passengers by Observation Techniques (SPOT) protocol[1] (Committee on Science, Space and Technology, 2011; Meyer, 2010; U.S. Government

[1] For comparison the United Kingdom does not follow the SPOT program and instead supports the development and implementation of behavioural detection methods based on scientific evidence (CPNI, 2020).

Accountability Office, 2010) and on the street with the use of the controversial stop and search policy ("stop and frisk" in the USA) used by police forces (Bradford & Tiratelli, 2019). However, the use of such measures relies on two main assumptions: that there are non-verbal behaviour markers that can identify those who may pose a threat; and that these assumed markers are easily and accurately observed in real-time with the naked eye.

The TSA's Screening of Passengers by Observation Techniques (SPOT) protocol is a behavioural detection program that was developed to screen for and thwart potential terrorist attacks at airports. The purpose of SPOT is to detect whether people are deceptive/pose a threat, thus requiring further questioning and possible law enforcement intervention. The SPOT program consists of three main stages (U.S. Government Accountability Office, 2010). First, behavioural detection officers (BDOs) visually assess those who are queueing prior to going through a security checkpoint. Second, having been trained to identify a series of behavioural indicators, the BDOs assess passengers for said indictors, which if present permits the officers to make an approach and engage in conversation with the passenger of interest. Third, passengers who display a certain number of cues are selected for further questioning or passed on to law enforcement. The SPOT program has received a lot of attention in recent years for its lack of scientific foundation, specifically regarding the validity of the training which makes BDOs look for specific discrepancies in non-verbal behaviour (U.S. Government Accountability Office, 2010; U.S. Government Accountability Office, 2017; ACLU, 2017). The thirty-six behavioural indicators (U.S. Government Accountability Office, 2017), sometimes reported to be as many as 92 (Winter & Currier, 2015), includes non-verbal behaviour such as yawning excessively, exhibiting confusion, improper attire, gazing downwards, eye contact with others, having a pale face due to recently shaving a beard and whistling, to name a few. These behavioural indicators are arguably either irrelevant to deception (Denault et al., 2020) or potentially prone to subjective interpretation.

Much of the critique of the TSA's SPOT program revolves around the use of invalid sources to justify the use of the non-verbal behavioural indicators that officers are trained to detect (U.S. Government Accountability Office, 2010). Worryingly, recent research has found that at least 70% of the behavioural indicators have no valid scientific evidence to support their inclusion in the SPOT program. Ninety-eight percent of the sources that the TSA cited in response to questions posed by the U.S. Congress, were invalid due to the use of opinions and poor quality or irrelevant research as justification of the SPOT program (U.S. Government Accountability Office, 2017). Amongst its list of grievances with the SPOT program, the American Civil Liberties Union (ACLU) states that the behavioural indicators that the BDOs are trained to detect seem to encourage profiling based on perceived racial and religious characteristics (ACLU, 2017). As a consequence, people perceived as Arab, South-Asian or Muslim can be disproportionately targeted by this security protocol (ACLU, 2017).

In addition to the criticisms of the use of non-verbal measures in the SPOT program, the academic scepticism of using non-verbal cues of deception revolves around their low predictive value, validity and reliability which varies between studies and individuals (DePaulo et al., 2003). Although research has found a preference for using non-verbal cues more than verbal cues (Bogaard et al., 2016) there is either no difference in accuracy when using non-verbal cues (Hicks & Ulvestad, 2011) or poorer accuracy rates compared to using verbal cues (Vrij, 2008b). Given that the non-verbal measures used in SPOT have dubious scientific grounding and within academia, the use of non-verbal measures is contentious, one of the core purposes of this thesis is to assess the suitability of non-verbal measures, specifically in high-security contexts.

Some important questions emerge given that non-verbal cues are being used in airport environments: is this valid given the current unreliability of the findings in the literature and the contextual relevance of research on deception?; are people able to distinguish between honest and deceptive people via passive observation in real-time or are people's judgements of the intent of others no different from guessing?; are people likely to rely on superficial characteristics to inform their decision-making, leading to discriminatory practices and an ineffective system of detecting threats?; and can we introduce de-biasing measures to reduce the incidence of relying on superficial characteristics and produce acceptable accuracy rates of detecting deception?

2.2 Displaying non-verbal cues of deception

With the ongoing popularity of crime television shows and films, in which the occurrence of deception and criminal behaviour is overrepresented for entertainment value (Dowler et al., 2006), it would be fair for people to assume that deception is easily identifiable from the observation of non-verbal behaviour. The popularity of crime television shows such as Lie To Me may contribute to and be reflective of false assumptions about the extent to which the truth can be easily distinguished from a lie (Levine et al., 2010). This assumption is evident amongst both layperson and expert (police, judiciary etc) populations (Jupe & Denault, 2019) and is rooted in the theoretical supposition that engaging in deception affects non-verbal behaviour in an observable and consistent manner regardless of the individual or environment.

2.2.1 The multi-factor theory's view of non-verbal deceptive behaviour

Zuckerman et al.'s (1981) multi-factor theory is one of the most prominent models in the deception field which provides an explanation for how and why people may exhibit cues of deception. At the crux of this theory is the assumption that behaviour is markedly different when people engage in deception compared to when they are honest. This theory is frequently cited in some form as an explanation for the behavioural differences between liars and truth-tellers found in studies of non-verbal behaviour and deception (Ekman, 1988; Eubanks et al., 2015; Vrij et al., 2017).

According to the multi-factor theory, engaging in deception is thought to affect cognitive load, emotional arousal, and attempts at behavioural control. Cognitive load refers to the finite resources that the brain has to process and store information and thus the limited capacity to engage in tasks, i.e. deception, that put the working memory under strain (Chandler & Sweller, 1991). Emotional arousal refers to being in a heightened state of psychological or physical arousal due to engaging in deception. Behavioural control refers to the extent to which one has control over the behaviours that they exhibit. Increased behavioural control denotes an attempt to override automatic behavioural expressions with more deliberate forms of expression (typically to form a good impression). In contrast, decreased behavioural control refers to an inability to influence which behaviours are displayed, potentially due to an increase in cognitive load or emotional arousal. The impact of the three factors may be visible through non-verbal behaviour and are often termed *non-verbal cues* (Zuckerman et al., 1981). Paramount to this theory is the view that honest and deceptive people are not influenced by the aforementioned factors to the same degree, thus theoretically making non-verbal measures a vital tool in security settings for detecting threats/deception.

The cognitive load and emotional arousal approaches, in particular, rely on the notion that the resultant non-verbal behaviours when engaging in deception are non-conscious automatic reactions to environmental pressures and stimuli (Depaulo et al., 1988; Eubanks et al., 2015) which are therefore difficult to conceal. The following will explain the rationale behind how deception may impact each of the three factors (cognitive load, emotional arousal and behavioural control), discuss the non-verbal findings in the literature and assess the extent to which the multi-factor theory is a suitable explanation of non-verbal deceptive behaviour.

2.2.1.1 *Cognitive load approach*

Cognitive load refers to the assumed strain of engaging in deception on working memory (Baddeley & Hitch, 1974; Walczyk et al., 2013). This presumed increase in cognitive load is reasoned to be the result of simultaneously constructing a believable lie, whilst inhibiting the truth (Sporer & Schwandt, 2007) and maintaining an internally consistent deceptive narrative (Zuckerman et al., 1981), which must also be consistent with external information (Granhag et al., 2015). All of which are posited to be more cognitively taxing than simply recalling the truth. Ultimately, it is theorised that being preoccupied with cognitively taxing verbal tasks, particularly during an interrogation, leads to changes in the non-verbal behaviour of deceptive people, such as a reduction in movement (Sporer & Schwandt, 2007; Vrij, 1995).

However, there may be an issue with generalising the assumptions of the cognitive load approach to an airport setting due to the predominantly seated interrogation methodology of a lot of the studies that have supported this theory. Whilst previous non-verbal findings may apply to interrogations, it is difficult to conclusively say the same for an airport environment. For one, non-verbal behaviour could differ when seated, owing to the reduced freedom of movement, compared to standing or walking around as one does in an airport (Vrij, 2008a).

The application of the general assumptions of the cognitive load approach may not be suitable without acknowledging the unique contextual pressure that high-security settings may also have on honest people; a contextual factor that may be lacking in laboratory environments. Perhaps the honest conditions of some interrogation based laboratory experiments do not exert an equivalent level of the contextual pressures that even honest people may feel in an airport environment (Blackwood et al., 2015). In airport environments, honest people may be under a similar level of strain as deceptive people, due to the goal of wanting to appear normal and be permitted to travel (Blackwood et al., 2015). Consequently, some of the non-verbal behavioural differences found in the literature may be specific to the context of those experiments. Therefore, it may be unwise to assume that we have a clear understanding of the applicability of the cognitive load approach regarding non-verbal behaviour in airport settings.

The cognitive load approach in experimental settings typically relies on the inclusion of an open-ended cognitive interview, in which cognitive load can be manipulated to observe differences in non-verbal behaviour between liars and truth-tellers. Within a laboratory setting, the largely unpredictable nature of the questions within the interview aid in increasing cognitive load. However, there are some issues with the transferability of this technique in the context of an airport. Realistically there are a limited set of plausible questions that may be asked about the purpose of travel and the content of their luggage/possessions. This means that it may not be feasible that passengers are under similar levels of cognitive strain in an airport environment as they are in laboratory settings. As a result of the potential discrepancy in the limits of cognitive load in a laboratory environment and an airport, some studies may have artificially inflated cognitive load to the extent that the observed non-verbal behavioural differences may not apply to behavioural responses in an airport.

Another reason why some of the cognitive load approaches may not be suitable for application in the aviation industry is that they rely on laborious open-ended interviews. The aviation industry is constantly assessing the cost-benefit ratio of introducing security interventions against the increases in the time it takes for passengers to pass through security because of the negative impact it could have on passenger satisfaction (Gkritza et al., 2006; Hainen et al., 2013; Omar & Miah, 2014; Stewart & Mueller, 2014). Consequently, applications of the cognitive load approach, using open-ended questions and techniques such as prompting reverse order recall (Vrij, Mann, et al., 2008), while potentially beneficial for distinguishing between honest and deceptive people in a laboratory, may not be practical for non-verbal assessment in an airport environment. Such techniques cannot be applied to all passengers in an airport environment due to time constraints and should not be applied due to the potential discrepancy in cognitive load effects in the laboratory compared to the airport.

2.2.1.2 *Emotional arousal approach*

Deception has been thought to lead to changes in non-verbal behaviour due to increases in emotional arousal (American Polygraph Association, 1997), which may be spurred by a fear of getting caught, feelings of guilt due to being deceptive or the excitement of the deception going undetected (Eubanks et al., 2015). The punishment theory (Davis, 1961), states that the prospect of the potential ramifications of a failed attempt to convey a deceptive narrative can lead to increased psychophysiological arousal and intensity of emotions (Ströfer et al., 2016). This arousal reflects the increased activation of the sympathetic nervous system (SNS). The SNS is the neural pathway that governs involuntary responses (Alshak & Das, 2020). Activation of the SNS, in response to a perceived source of stress, can lead to an increase in adrenaline and cortisol to prepare for the fight or flight response (Allison et al., 2012; American Psychological Association, 2019; Ayada et al., 2015). Deception induced activation of the SNS can result in responses such as increases in heart rate, perspiration and pupil dilation (Ziegler, 2004). The potential impact of deception on emotional arousal spurred the prominence of the polygraph as a method of detecting deception from the 1920s (Grubin & Madsen, 2007). The polygraph is still commonly used in South Africa (Nortje & Tredoux, 2019), as well as in the USA and the UK (Home Office, 2020; Marshall & Thomas, 2015), where it is increasingly used to assess sex offenders on probation. The polygraph indirectly measures emotional arousal by detecting physiological changes, such as respiration, electrodermal activity and blood pressure, which may be interpreted as an indication of deception (Meijer & Verschuere, 2010). Though this test, much like all deception detection tools, cannot measure deception directly, changes in the physiological factors are often used to indirectly infer whether someone is deceptive.

Still, emotional arousal, spurred by fear or excitement, in an airport context is largely understandable. It is conceivable that excitement around travelling, fear due to the prospect of flying or an armed security presence are not abnormal responses (Blackwood et al., 2015). Feelings of guilt when being deceptive should not be assumed because it may be mitigated by one's own beliefs about the justification for acting deceptively. For example, the *Spiritual Manual* that the 9/11 attackers used suggests that the attackers possessed little guilt owing to the perceived "pious deed" they were to commit (Kippenberg & Seidensticker, 2006, p. 25) since jihad and other forms of "holy wars" by definition are interpreted by those who pursue them as being justifiable. Ultimately, the expression of fear, guilt and excitement is not guaranteed in all contexts, nor for all people. As a result, interpreting the cause of non-verbal cues attributed to emotional arousal, such as micro-expressions (Jupe & Keatley, 2019) is difficult and casts doubt on the validity of the passive observation of non-verbal cues given that in an airport environment, emotional arousal may occur for innocuous reasons.

2.2.1.3 *Behavioural control approach*

Engaging in deception is also theorised to increase the likelihood that people attempt to control aspects of their behaviour to increase the chances of their deception being successful.

Two factors may contribute to attempts at behavioural control. For one, intentional attempts to control behaviour may be directly influenced by an individuals' beliefs about which behaviours indicate deception (Vrij et al., 1996; Zhang et al., 2013). This approach suggests that increased behavioural control is a voluntary attempt to modify perceived involuntary expressions of deceptive non-verbal behaviours. To illustrate, believing that liars can be identified from their hand movements was found to result in a distinct reduction of hand and finger movements (Vrij, 1995). Increased stiffness and a lack of spontaneity in physical movements (Vrij et al., 1996) are also often cited as evidence of people employing methods of behavioural control when they are being deceptive. These attempts to control behaviour then become cues of deception in themselves, possibly because they are informed by invalid beliefs, making their modified behaviour appear out of place or stiff, rather than reflecting honesty.

The attribution of non-verbal cues to behavioural control is one example that highlights the unreliability of the display of non-verbal cues and their questionable inclusion in security protocols. Contradicting with the assumed increase in behavioural control techniques when engaging in deception, several studies have surmised that engaging in deception can cause an unintentionally reduced capability to control behaviour, due to the impact of cognitive load and psychophysiological arousal (Vrij et al., 1996). Limited cognitive resources during deception are thought to result in behavioural leakage (Ekman & Friesen, 1969) which refers to an increase in behaviours that may indicate deception due to a reduced capacity to control which behaviours are displayed. (Ekman et al., 1976). An example of behavioural leakage is the display of gestural slips (Cohen et al., 2010), which are characterised as hand or head gestures that semantically contradict the corresponding verbal communication. Zloteanu (2016) made an important distinction between *authentic cues*, which are involuntary, similar to Ekman's leakage theory (reflecting reduced behavioural control), and *inauthentic cues* which are deliberate attempts at increasing behavioural control. There is a danger in the use of non-verbal cues which can be attributed to behaviour control (and hence indicative of deception) given that the literature suggests that it may lead to both diminished movement or the presence of unintended movements (e.g. gestural slips). Consequently, an unjustifiably wide range of behaviours could be treated with suspicion, most of which may be innocuous.

2.2.2 The self-presentation theory's view of non-verbal deceptive behaviour

Despite all of the mentioned approaches on the display of non-verbal cues of deception, DePaulo et al. (2003) found that very few non-verbal cues were reliable or were strongly correlated with deception. The DePaulo et al. (2003) meta-analysis is often cited in deception research, but despite it being almost two decades old at present, more recent research suggests that little has changed since then (Luke, 2019). The multi-factor model may provide an ideal explanation of non-verbal behaviour in some contexts but if Zuckerman et al.'s (1981) theory were all-encompassing, the non-verbal cues in the literature would not be as unreliable and oftentimes contradictory. n alternative to the multi-factor model is DePaulo's (1992) Self-Presentation theory. The crux of this theory is in contention with the cognitive load and emotional arousal approaches of the multi-factor model, because it proposes that deception related non-verbal behaviours are not as involuntary as assumed. Rather the self-presentation theory suggests that non-verbal behavioural changes are solely the result of purposefully modifying behaviour solely to convey a good impression. Importantly, the self-presentation theory differs slightly from the behavioural control approach of the multi-factor model in that it suggests that both honest and deceptive people modify their behaviour in some contexts to be perceived positively, whereas the multi-factor model would suggest that only deceptive people modify their behaviour. Subsequently, one particularly important aspect of the self-presentation theory is that it highlights that there are more similarities than differences between liars and truth-tellers' behaviour because both groups wish to portray a good image of themselves (Granhag et al., 2015). Both groups of people, according to the self-presentation theory, aim to portray themselves in the best light and may modify their behaviour to the extent that they think they appear harmless. Accordingly, there may be little difference in the non-verbal behaviour of liars and truth-tellers.

The self-presentation theory may be closer to reality than the multi-factor model, especially in the post-9/11 heightened security environments that we face today. Some sub-sections of society, for example, minority groups who may be more at risk of being profiled, such as Muslims, have reported modifying their behaviour, despite their harmless intentions, in an attempt to avoid arousing suspicion at airports (Blackwood et al., 2015). This purposeful modification of behaviour even when honest contrasts with the *illusion of transparency* which presents the idea that people overestimate the extent to which their inner state is perceptible to others (Gilovich et al., 1998). This illusion extends to people's beliefs that if they are innocent, others will be able to perceive this (Mandelbaum, 2014). Opposingly, Blackwood et al.'s, (2015) reporting of some Muslims' behaviour modification at airports may suggest that the illusion of transparency is mediated by factors such as belonging to a religion, ethnicity or race that is actually or perceived to be disproportionately targeted in particular contexts (Choudhury & Fenwick, 2011; Lewis & Marsden, 2020).

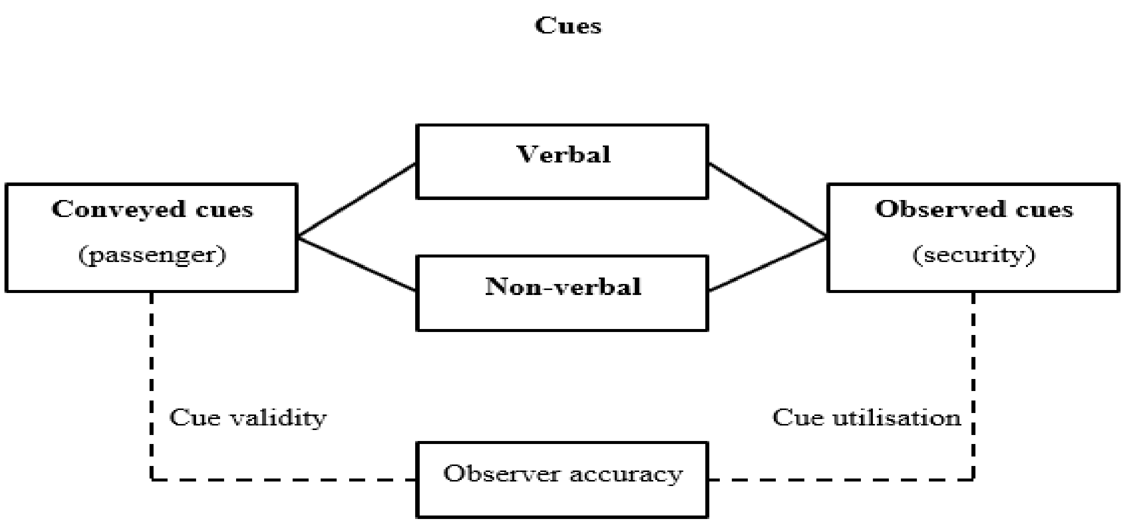
Accordingly, the self-presentation theory has some important ramifications for our understanding of non-verbal behaviour in high-security environments and the design of security protocols. Namely that the context of high-security environments and the desire not to draw the attention of security personnel may not solely impact the behaviour of deceptive people, potentially to the extent that there are no stable overall differences compared to honest people. Consequently, the lack of consensus in terms of the presence of significant cues, the direction of effects, validity and reliability across studies may be due to flawed theoretical assumptions. Some of the non-verbal findings in the literature could be mere artefacts of the traditional laboratory environments/paradigms in which they were gained, and not reflective of behaviour in high-security environments. This conflict between the multi-factor approach and the self-presentation theory informs part of the motivation to investigate non-verbal behaviour in this thesis, specifically in airport contexts. The method of simulating a high-security environment to assess differences in honest and deceptive non-verbal behaviour will be expanded on in section 2.4.1.

2.3 Identifying non-verbal cues of deception

The previous section focused on the exhibition of non-verbal cues whilst this section will focus on the process of identifying deception. First, Brunswik's lens model (Brunswik, 1952) will be explained because its approach informed the structure of this thesis. Second, four models will be discussed (the signal detection theory, the expectancy violations theory (Burgoon et al., 2005) and the motor and perceptual experience hypotheses (Cañal-Bruland et al., 2010)) which explain the overarching paradigms and explanations for deception detection capabilities in laboratory settings and the field. Third, some studies on people's accuracy of identifying deception will be discussed.

A Brunswik's lens model approach (Brunswik, 1952) can be useful in deception detection research (Hartwig & Bond, 2011; Sporer et al., 2014). The lens model is a conceptual framework for decision-making, which denotes decisions as a probabilistic function of both cue use and cue validity in situations where there is an objectively correct decision. According to the model, decision-making accuracy is related to the presence and appropriate use of deception cues, as well as the cues having diagnostic power. In the context of a security checkpoint, this model ascribes importance to the validity of the relationship between the non-verbal behaviours that passengers exhibit and deception, as well as whether and how these cues are perceived by security personnel (see Figure 2.1). Consequently, this framework allows us to ascertain whether poor accuracy is due to a lack of diagnostic validity of the cues, poor use and interpretation of cues to inform decision-making, or both (Hartwig & Bond, 2011). The lens framework is particularly useful within deception research to conclude whether non-verbal cues are suitable for deception detection. As such, this thesis has been structured so that measures of deception are investigated in Chapter Four (cue validity), and the accuracy of detecting deception and the factors that influence people's judgements are assessed in Chapters Five to Seven.

Figure 2.1 A Brunswik's lens model of deception detection.



Note. Adapted from Brunswik, E. (1952). The conceptual framework of psychology, Journal of Consulting Psychology

2.3.1 Theoretical approaches to identifying deception

The following will discuss some theories regarding the methods that people may use when attempting to detect deception in others. The signal detection theory (SDT) assumes that there is a fixed threshold for defining normal behaviour when making decisions where there is a degree of uncertainty (Green & Swets, 1966). This method is reflected in many security protocols globally wherein security personnel are trained to assess a specific set of verbal and/or non-verbal behaviour for signs of deception (Wigginton et al., 2014). This practice is akin to using the SPOT program’s behavioural indicator checklist, the use of which culminates in subjecting passengers who exhibit a particular amount of behavioural indicators to more intense questioning or detainment (Winter & Currier, 2015). One of the major issues with signal detection-based methods is that they may increase the chance of false negatives because people who do not exhibit the pre-determined cues, despite posing a threat, may go undetected. Another important negative consequence of this method in security protocols is an increased chance of false positives may arise as people may innocently display some of the cues, such as SPOT’s inclusion of “exaggerated yawning” which likely reflects tiredness, not a terroristic threat (ACLU, 2017, p.8).

In contrast, the expectancy violations theory (EVT) (Burgoon et al., 2005) proposes a method of deception detection without a pre-determined set of cues, unlike SDT. Rather, any behaviours that violate what would typically constitute normality in that environment result in one of two violations. A positive violation, such as someone standing close to security personnel, could enhance perceptions of trustworthiness because it violates the expectation that someone with something to hide would try to keep their distance (Burgoon, 2015). A negative violation, such as failing to correctly answer an obvious question about plans for a trip, may result in being judged with suspicion. This theory implies that there is a baseline of normal behaviour in each context from which discrepancies can be ascertained. When used in the context of a security checkpoint, EVT may prove difficult to incorporate in real-time. Many innocent variables could influence one's behaviour, leading to a violation of expectation and causing suspicion to arise. For example, the increased security presence at major airports and nervousness about flying can cause anxiety in some people (Busscher et al., 2013), resulting in behaviours perceived as violating the norm. Also, training staff to become proficient at EVT methods with passengers from a range of cultures, ethnicities and religions is a difficult task that could be susceptible to subjective decision-making, informed by prejudice (Meyer, 2010). Behaviours which violate the norms of the mainstream culture may be perfectly normal within specific sub-cultures (Vrij & Winkel, 1991), but could trigger a violation of expectation in culturally unaware staff, regardless of a baseline having been established. Therefore, there are challenges in applying EVT to a security context even with staff training.

In practice, both EVT and SDT may be used to a degree. For example, in the SPOT program, the essence of EVT motivates the initial arousal of suspicion and then an SDT-like checklist is used to further assess the person of interest (Committee on Science, Space and Technology, 2011) . Both models of deception detection may increase the likelihood of security personnel being influenced by biased reasoning because they leave the onus on security personnel to decide if the non-verbal behaviour is suspicious. The potential for preconceptions to influence decision-making has inspired part of this thesis which introduces technology to reduce the salience of superficial characteristics and assess the effect of information reduction on decision-making accuracy. This will be examined in Chapter Six and briefly mentioned in section 2.4.2 in the discussion of automated methods having the capability to help debias decisions.

2.3.1.1 *Insights from the sports domain*

Outside of the forensic context, other theories of deception detection have been put forward to explain what contributes to achieving good accuracy at identifying deception. The motor and perceptual experience hypotheses (Cañal-Bruland et al., 2010) are two similar concepts about how people may detect deception, which have been largely derived from sports psychology. The motor experience hypothesis attributes better accuracy at detecting deception to increased personal experience of performing deceptive acts. This hypothesis is in agreement with our knowledge of motor simulation, more specifically the Action Observation Network, which specifies that observing the actions of another person can activate areas of the motor cortex in the observer (Grafton, 2009; Iacoboni et al., 2005). Accordingly, if the same areas of the brain are activated when one is committing a non-verbal deceptive act of which they have experience and when watching someone else commit a deceptive act, one should be better at identifying the deceptive act than someone with less motor experience. This explanation would explain the better accuracy of expert badminton players compared to novices when identifying deception in other players (Park et al., 2019). Whilst the literature from sports psychology provides an interesting insight into the impact of motor experience on deception generally, the type of deception in these studies is very different from those of interest in security-related deception detection. Furthermore, the practical transferability of the motor hypothesis is limited in security-related deception detection due to the implication that observers, i.e., security personnel, require motor experience of being deceptive to gain from this explanation of deception detection.

In contrast, the perceptual experience hypothesis (Cañal-Bruland et al., 2010) attributes better accuracy at detecting deception to the increased visual familiarity of seeing others behave deceptively, independent of motor experience. The improved accuracy is related to the experts' increased amount of exposure to seeing other peoples' deceptive moves. The motor and perceptual hypotheses are somewhat intertwined because increases in motor experiences of deception are likely related to increases in perceptual familiarity. For example, some populations such as prisoners may have more motor experience of acting deceptively and potentially more perceptual experience from viewing others being deceptive within the prison population. However, increased perceptual familiarity does not necessarily always coincide with increased motor experience. For example, whilst security guards are more likely to witness deception in their line of work than laypeople, they may not engage in deceptive behaviour themselves. Importantly, expert populations, on the whole, do not typically show superior accuracy compared to lay populations (Bond, 2008), suggesting that the nature and context of deception in sports research may limit the generalisability of these hypotheses beyond the sports domain. Whilst both the motor and perceptual experience hypotheses provide an interesting insight into how accuracy can be improved, they are of little value to improving deception detection within security contexts.

This section has elaborated on how Brunswik's lens model has been used as a structural framework in this thesis. The lens framework ensures that non-verbal cues are assessed in Chapter Four, and people's accuracy at detecting deception are assessed in Chapters Five to Seven to make an overall conclusion about non-verbal behaviour and deception. The theories which may be relevant to how people form their decisions of whether others are deceptive were also discussed, highlighting their practicality and the detrimental potential for subjectivity to influence decision-making using SDT and EVT.

2.3.2 Poor accuracy detecting deception

Despite the number of non-verbal cues of detecting deception that have been investigated in the literature, people tend to perform poorly, an average of 54% accuracy, when identifying whether others are deceptive (Bond & DePaulo, 2006). This is worrying because it implies that people's ability to detect deception is little different than guessing, bringing into question whether non-verbal assessments should be included in security protocols.

There are several reasons why accuracy may be so poor. At first glance, a fair critique of the low accuracy rates in the literature revolves around the predominantly student population as opposed to the experts that are tasked with this role in the real world. Yet only a limited number of studies have produced good accuracy rates, above 80 %, even with expert populations. Bond (2008) found that accuracy rates among experts ranged from 31.25 % to 93.75%, accounting for a large amount of variability amongst this expert population which consisted of local and federal law enforcement in the USA. However, only 10% (11 participants) of Bond's sample had accuracy rates above 80%, exemplifying an issue with the few studies that report impressive accuracy figures amongst experts. The deception detection savants reflect a very small proportion of the expert population, indicating that security personnel as a whole may not be good at identifying deception. Average accuracy rates did not differ significantly between student and expert populations (Bond, 2008), suggesting that the predominantly student participant population is not a detriment to our understanding of detection abilities and may not be driving the low accuracy in the literature.

A few other factors may contribute to poor deception detection. These factors include the idiosyncratic nature of non-verbal cues; issues with the approaches to perceiving non-verbal behavioural differences if there are any; decisions may be biased by the salience of superficial factors; people may use invalid cues informed by pseudoscience/misconceptions. The latter is an issue that can affect the accuracy of individuals tasked with identifying deception as well as the incorporation of pseudoscientific information into security protocols.

2.3.2.1 *The impact of pseudoscience on accuracy*

One of the prevailing ideas about the low accuracy rates is that people may rely on invalid cues to inform their judgements of deception in others (Stel et al., 2020; Strömwall & Granhag, 2003; Vrij & Semin, 1996). The use of invalid cues may be influenced by perceptions of deception cues as seen in popular media, which can hurt deception detection accuracy (Hayes & Levett, 2013). A reliance on invalid cues due to possessing inaccurate beliefs is evident in both layperson and expert populations (Bogaard et al., 2016; Denault, 2020). The pervasiveness of inaccurate beliefs presents a potential risk of being implemented in the real world, such as by detectives within interrogations (Strömwall & Granhag, 2003) and through security protocols at airports with the possibility for disastrous ramifications. Though the studies cited in this thesis are largely limited to Western countries, false beliefs about deception may be universal. A worldwide cross-cultural study, which included 75 countries and 43 language groups, found that the belief that liars avoid eye contact was universal, despite the literature tending to show that there are no differences in eye contact when telling the truth or lying (The Global Deception Research Team, 2006). Though the layperson sample of this study means that the ramifications of these misconceptions are not as critical an issue compared to if those beliefs were held by experts, worryingly, the use of pseudoscience (invalid measures of deception) is also an issue within law enforcement (Lilienfeld and Landfield, 2008), which could lead to poor accuracy whereby irrelevant factors are relied on to identify deception. The prominent use of protocols such as the Reid technique in police interrogations (Inbau et al., 2001), which was developed largely without any scientific backing and has been criticised for its coercive nature (Kozinski, 2018) potentially reflects the spread of false beliefs on an institutional level as well as amongst individuals. This demonstrates how pseudoscience can seep into real-world procedures.

Police often report having a “sixth sense” to justify their perceptions of others’ intentions (Pinizzotto et al., 2004, p. 6) but this could simply signify the potential for superficial characteristics to bias decision-making. The use of unscientific methods of detecting deception has the potential to allow discrimination to either consciously or sub-consciously impact security processes given the potential for irrelevant factors such as perceived unattractiveness (Murai et al., 2018), scruffy clothing and reduced eye contact (Vrij, 2000) to increase perceptions of deceit. Decisions influenced by these superficial characteristics could lead to increased and unjustified scrutiny of certain sub-sections of the population (Herbert, 2007).

In truth, the majority of people tell relatively few lies daily, if any (Depaulo et al., 1996; Serota et al., 2010). The truth bias or truth-default theory proposes that people are pre-disposed to believe that the majority of communication is truthful (Levine, 2014). This default could be an evolutionary adaptation that encourages cooperation amongst people. However, the truth bias is not apparent amongst all populations. The opposite is evident in police officers, as they are more likely to assume that people are lying compared to non-officers, (Masip & Herrero, 2017; Serota & Levine, 2015). The lie bias can be detrimental for the people whose behaviour is being judged due to the potential for misplaced suspicion. These false beliefs about the pervasiveness of deception throughout society can also permeate into security protocols and personnel. The prevalence of the lie bias amongst expert populations could have disastrous effects given how things can spiral from an initial misclassification of deceit to a false confession and potential wrongful conviction (Innocence Project, 2019; Leo, 2009).

The aforementioned findings cast a bleak look on the universality of false beliefs about deception, nevertheless, some populations seem less susceptible than others. Vrij and Semin (1996) found that prisoners had more accurate beliefs about the indicators of deception compared to students and an expert population (police, detectives, prison guards and customs officers). These findings may be in line with the motor experience hypothesis (Cañal-Bruland et al., 2010), suggesting that prisoners' increased likelihood of deceptive motor experience may have contributed to more accurate beliefs. Whilst acknowledging the contribution to knowledge that Vrij and Semin's (1996) study provides in terms of the variation in inaccurate beliefs, the practical impact on the security sector is limited. There is little transferability of these findings to potentially remedy the deception detection accuracy of experts and others without this motor experience. Instead, the need to assess whether non-verbal cues are appropriate in security protocols dictates that research focus on what influences non-prisoner populations' deception detection and methods that could aid their performance.

In contrast, Hartwig & Bond (2011) dispute the prevalent idea that people have invalid beliefs about deception which contributes to poor accuracy. Instead, they suggest that while people's explicit self-reported perceptions of deception cues may be inaccurate, intuitively, they are accurate at pinpointing valid cues. According to Hartwig and Bond (2011), poor accuracy at detecting deception may be improved by manipulating the decision task so that intuitive skills are utilised more, rather than training people to look for specific cues of deception outright. If Hartwig and Bond's hypothesis is correct, then introducing techniques such as time constraints

(Albrechtsen et al., 2009) could result in more acceptable accuracy levels. However, their hypothesis somewhat relies on the multi-factor theory's notion that there are valid differences in behaviour, which then can be used to inform judgements. If the self-presentational theory is more accurate and there are few differences in behaviour, then according to the lens theory (Brunswik, 1952), poor accuracy is merely a reflection of a lack of predictive cues and is therefore inevitable.

2.3.3 The research-practice gap

Regardless of whether people may only consciously have invalid beliefs, the incorporation of invalid non-verbal assessments in security protocols could be remedied by reducing the "research-practice gap" (Brownson et al., 2018, p.102), which refers to the poor dissemination of research findings beyond academic circles. This dissemination issue is not limited to deception research, nor psychology, but can be attributed to academia as a whole. Jupe and Denault (2019) suggest that the pervasiveness of false beliefs, which also extends to prosecutors and judges, may be due to misunderstandings over what distinguishes science from pseudoscience. This finding may be compounded by the relatively poor distribution of research beyond the academic community.

There are two aspects of the security community for which reliance on pseudoscience is problematic: the decision-makers informing policy and the security personnel carrying out the policies. The inclusion of non-verbal cues despite the lack of justification may be due to inconsistencies within the literature which may make it difficult for people to determine when and how to apply research findings to decision-making and policy design. The inconsistencies in non-verbal findings may be partially driven by the methodological differences between studies, which results in different levels of contextual pressure and conflicting findings. For example, in a study on verbal cues, Mapala et al (2017) used virtual reality to simulate an airport and asked questions directly relevant to a security context. They found that response latencies were shorter when deceptive compared to when honest, conflicting with the literature (Sporer & Schwandt, 2006). Consequently, although studies may report valid cues for detecting deception, the applicability of these cues beyond the context of the experiment is uncertain as cited by practitioners who may have doubts about the usefulness of laboratory studies to security settings (Buckley, 2012). This lack of generalisability may contribute to the use of invalid cues in real-world security procedures which are often designed with information that is gained from contextually relevant but anecdotal evidence from law enforcement (Committee on Science, Space and Technology, 2011). Establishing contextually relevant paradigms may lead to greater confidence in the relevance of the laboratory findings to the real world. Reducing the salience of laboratory environments may induce behavioural responses similar to a security context. This methodological adjustment may lessen the reliance on other pseudoscientific claims if the applicability of our research is more direct.

Another factor impacting the proliferation of scientifically dubious measures is their relative ease of application and perceived utility over scientific methods (Denault et al., 2020). For example, in the wider forensic context, the Reid technique (Inbau et al., 2001) is used by interrogators worldwide (Cleary & Warner, 2016) despite its propensity to induce false confessions (Kozinski, 2018) and its use of unsupported verbal and non-verbal cues (Denault et al., 2020). The active encouragement to use the Reid technique within law enforcement, despite the widespread condemnation of its lack of validity and propensity to lead to false confessions conveys how the pervasiveness of false beliefs can be systemic throughout the security sector. Likewise, the SPOT program was partially based on the anecdotal beliefs of law enforcement personnel and agencies due to their encounters with deceptive people (Committee on Science, Space and Technology, 2011). Proponents of pseudoscientific methods of detecting deception can make big claims about their effectiveness compared to scientific methods which come with a more realistic and conservative efficacy. This lack of efficacy can make them less attractive to those unaware of the importance of scientific validity.

The issue of the widespread use of invalid cues in security protocols brings to the fore the potentially inadequate level of communication between academia and those working within security or legal organisations. Jupe and Denault (2019) rightly point out that the use of invalid measures within organisations such as the TSA is not solely the fault of the organisations themselves. It reflects a lack of access or understanding of the research that academia disperses. Existing within these organisations is an apparent lack of knowledge of what constitutes scientific evidence and what is pseudoscience. Given that the validity of the SPOT program was called into question as early as 2010 (U.S. Government Accountability Office, 2010), there seems to be a lack of understanding of the harm that is posed by invalid measures (Denault & Jupe, 2018). Greater collaboration between academia and policymakers is one such method of reducing the likelihood of irrelevant or unsubstantiated information informing policy.

Ultimately there is a greater need for more interaction between academia and practitioners to help reduce the use of pseudoscientific methods. Studies that investigate cues of deception should incorporate more contextual relevance into their designs which may aid dissemination beyond academic circles and prevent the use of pseudoscientific anecdotal but contextually relevant methods.

2.4 Making advances in non-verbal deception detection with technology

Some of the issues that have been discussed thus far revolve around the importance of context in the behavioural outcomes of research within the deception field and consequentially, the validity/applicability of the knowledge gained to high-security environments. The potential for subjective and possibly discriminatory practices to influence decision-making, and the difficulty of achieving accurate judgements of deception has also been made apparent. This section will discuss how technology has the potential to help address these issues.

2.4.1 Contextually salient research methods

Some deception researchers emphasise the importance of high stakes in deception research (Levine, 2018; O'Sullivan et al., 2009; Vrij et al., 2010). Attempts to induce high stakes can vary, and include but are not limited to offering cash incentives, or emphasising that participants' performance may directly impact another person. However, research ethics prevents the incorporation of real high stakes. Participants are likely aware that they have signed up for a study in which the researchers are ethically obliged to mitigate any physical and emotional harm. As a result, the extent to which the artificial manipulation of high stakes in a laboratory is perceived to be reflective of real life is uncertain.

An alternative solution to attempt to maximise the relevance of laboratory findings for application in security contexts could include methods of increasing the contextual relevance of research. Contextual relevance needs to take precedence (Vrij et al., 2019) and is a factor that must be taken into consideration both for assessing the non-verbal cues that people display when deceptive, but also for assessing how accurate people are at detecting deception in others (Vrij et al., 2015). One way of manipulating the contextual relevance of the testing environment is by using virtual reality to simulate high-security scenarios such as airports. The increasing realism of virtual reality technology (de Groot et al., 2020) may help to reduce the salience of laboratory settings and may also induce behaviours that are more in line with the real world. Simulating an airport with virtual reality is a form of contextual priming (Cerda et al., 2021), whereby the familiarity of the airport context could trigger situationally valid behavioural responses. However, the use of contextual priming via virtual reality in non-verbal deception research is in its infancy. The rationale behind contextual priming partially stems from the use of context reinstatement to bolster memory recall, particularly within eyewitness identification research (Bailenson et al., 2008; Norman et al., 2020). The increasing use of virtual reality within the behavioural sciences builds on this notion of context reinstatement, suggesting that increasing the contextual relevance of the testing environment via virtual simulation could prompt more externally valid expressions of behaviour (de Groot et al., 2020). Therefore, the use of contextual priming via virtual reality in this research is largely exploratory to see what insights about non-verbal behaviour can be gained, in comparison to the information that is available from more traditional methods. Context can have an important impact on psychological findings (Reynolds & Rendle-Short, 2011) and accordingly, using virtual reality may make the findings clearer in terms of their applicability to security settings (Levine, 2018), and our understanding of non-verbal behavioural changes in these environments.

Virtual reality is fairly new in this specific area of deception research, so few studies have compared deceptive behaviour in virtual reality, a traditional laboratory setting and the real world. However, it would stand to reason that there is a benefit of reducing the salience of laboratory environments whilst maintaining control (Parsons, 2015), from a theoretical perspective, as a way of gaining more information about non-verbal deception. In an unrelated domain, Bhagavathula et al. (2018) found that participants' pedestrian behaviour did not differ in virtual reality and real life, which is promising for the utility of virtual reality in terms of transferability of findings, but no strong conclusions can be made given the difference in context. Contextual priming in virtual reality could be beneficial due to the flexibility of creating the environment that participants are immersed in, especially when the environment of interest is an airport, which is otherwise difficult to gain access to or replicate in a laboratory. Virtual environments can help to heighten the contextual relevancy of participants' behaviour by including relevant audio and visual information (Pollard et al., 2020), as well as external tactile stimuli to immerse participants in the virtual environment and make them feel as though they have been transported to a security environment as opposed to being very aware of the laboratory and researcher. Furthermore, given the replication crisis within the social sciences, which also affects the deception detection literature (Lilienfeld, 2017; Maxwell et al., 2015; Shrout & Rodgers, 2018), virtual reality can allow for near-identical methodologies between different research groups to better assess whether research findings are reliable across participants and cultures.

2.4.2 Manual versus automated methods of measuring non-verbal behaviour

Judging individual non-verbal cues has the potential to be used subjectively in terms of the consistency with which targets displaying such cues are assessed as deceptive due to the difficulty of detection given the multitude of causes for discrepancies in behaviour. In contrast to searching for an individual cue of deception, coined the *Pinocchio's nose* of deception, whole body gait analysis is an emerging aspect of non-verbal deception detection. Whole-body gait analysis considers the interconnected nature of physical movement and the benefit of a more gestalt approach that values the body as a whole rather than focusing on individual areas of the body for cues. Gestalt theory in its simplest form stresses that in some contexts perceptions of the whole are not always perceptible from their individual characteristics (Wertheimer & Riezler, 1944). Applying this theory to non-verbal cues of deception warrants a shift from researching individual cues and using methods that can incorporate a more holistic assessment of bodily changes associated with deception.

Alongside, the potential of focusing on the body as a whole, the fidelity of measurement is another issue that may be detrimentally impacting non-verbal cues in the literature. Luke (2019, p.3) highlights that the “unusually high flexibility in coding” within the deception cue field could potentially cause false positives in the literature. Quantifying changes in behaviour via human observation could leave the coding aspect of non-verbal behaviour research subject to bias. One way of mitigating this is via the use of tools that can measure changes in non-verbal behaviour and bypass the need for manual coding based on naked-eye observations of non-verbal behaviour. Instead of manual coding, automated methods of recording and quantifying behavioural changes may be a step in the right direction.

The use of technology has the potential to help de-bias the detection of deception, both in terms of the observation of non-verbal behaviour in the real world and researcher behaviour when quantifying behavioural changes. Previous sections of this literature review chapter have detailed the subjective nature of deception detection in practice, which can have ramifications for public safety and people’s treatment at airports. Reducing the subjectivity of decision-making in high-security contexts is important due to the potential for it to influence accuracy and lead to less discriminatory practices (ACLU, 2017; Meyer, 2010) As previously discussed, poor accuracy rates of detecting deception (Bond & DePaulo, 2006) may either reflect an inability to detect differences in behaviour via passive observation or suggest a lack of predictive cues. If the former, then automated measurement of non-verbal deception detection may be a more fruitful avenue given the supposed imperceptibility of non-verbal deception to the naked eye. The automated measurement of body analysis (Poppe et al., 2014) is one such tool that could help with debiasing research findings (discussed further in Chapter Three, section 3.1). Used in conjunction with Xsens™ motion capture suits, Van Der Zee et al. (2015) found that joint displacement was a good measure to distinguish between deceptive and honest people. Van Der Zee et al’s (2015) research emphasises the utility of motion capture in non-verbal research and provides support to the idea that deception research should look into shifting from manual coding of non-verbal cues of deception to more precise, technology-based methods to lessen possible researcher influence on the outcome (Luke, 2019). As well as trying to reduce the influence of researchers in the coding stage, tools such as this allow for a more gestalt point of view, wherein the combination of bodily changes can be holistically assessed, as opposed to just looking at individual cues which have shown little predictive value so far (DePaulo et al., 2003). However, the contextual setting of the aforementioned studies using automated methods makes it difficult to ascertain whether such differences would also be apparent in dynamic non-verbal measures reflective of behaviour in an airport context, as opposed to the seated interrogation style environment that they used.

Poppe et al. (2014) suggest that automated-measurement methods of detecting deception, using non-verbal cues, may be better than manual methods because they can overcome some of the factors that may be hampering detection currently. Prejudicial judgements based on superficial characteristics such as race, gender and ethnicity (ACLU, 2017; Hashad, 2004; Winter & Currier, 2015), as well as invalid beliefs about deception are some examples of the factors which may be detrimental to detection and which automated methods could remedy by reducing the potential for homing in on superficial or irrelevant behaviour. Poppe et al. (2014) found that an automated data mining method was more accurate at detecting deception than human-based measures. The context of Poppe et al.'s (2014) research, consisting of a seated interview in which participants were instructed to steal money from a wallet, is very different from the types of deception you may expect in an airport environment. Nonetheless, it provides an important insight into both the potential for small changes in behaviour to be scientifically measured with motion-capture and later applied to machine learning methods of detecting deception.

Realistically, security personnel can assess fewer aspects of behaviour at a time, unlike automated measurement methods which can detect deception based on a multitude of non-verbal cues, taking into account the precise way that these cues manifest. Automated methods of detecting deception may be superior to human observations because they can distinguish between small, potentially unobservable differences (Bartlett et al., 2014; Lu et al., 2005; Michael et al., 2010). For example, if there are general differences in non-verbal behaviour between honest people and those with deceptive intent even when not actively lying, this could be used with an automated CCTV gait detection tool (Bouchrika et al., 2011). This could be a time-effective method in the security assessment process, since it would detect deception even in the absence of vocalised deception, purely due to having deceptive intent. Poppe et al.'s (2014) findings, along with the controversy regarding the SPOT program’s use of subjective behavioural indicators, bolsters the idea that automated-measurement methods of detection could need to have a bigger role in the real world instead of human-observation based diagnoses of deception. Should this thesis find that automated-measurement methods of measuring behavioural differences reveal differences between liars and truth-tellers, this may lend credence to the development of more hands-off, real-time tools to identify deception (Bouchrika et al., 2011; Lu et al., 2005), incorporating the data gained from automated-measurement tools with machine learning. However, as Jupe and Keatley (2019) rightly point out, methods of deception detection based on flawed science simply exacerbate the problem. Yes, precise methods of detecting changes in non-verbal behaviour lend themselves to eventual automated, artificial intelligence tools, but automation in real-world contexts should not be a concern for the near future given our current understanding of the inconsistency of non-verbal cues of deception and the need for greater theoretical consensus.

2.5 Current thesis

This literature review has highlighted that passive non-verbal assessments are often used in the first steps of some behavioural detection programs, informing whether security personnel decide to approach people for questioning. However, the current state of the literature suggests that non-verbal cues of deception are unreliable, potentially influenced by context and decision-making is marred by invalid beliefs about deception. Furthermore, although Zuckerman et al.'s (1981) multi-factor theory underpins much of the search for non-verbal cues of deception and their inclusion in security protocols, the self-presentational theory (DePaulo, 1992) could be closer to the truth in airport contexts, deeming the non-verbal deception detection measures in practice wholly inappropriate. The context of studies of non-verbal cues of deception has often been geared towards seated police interrogation settings (Granhag & Strömwall, 2002; Vrij, 2008b; Vrij et al., 1996), as opposed to the airport environments of interest in this thesis. This contextual difference can make it difficult to generalise the findings, due to the potential for non-verbal behaviour to differ depending on context (Vrij et al., 2019). To address the issue of contextual relevance, this thesis used virtual reality as a contextual primer to try to reduce the salience of the laboratory environment and induce behaviours reflective of an airport. This thesis assessed: whether there were deception related differences in non-verbal behaviour; how accurate people were at detecting deception from non-verbal behaviour; what factors people used to influence their decision-making; methods of reducing reliance on superficial factors and what people looked at before deciding whether others are deceptive. Ultimately, the appropriateness of non-verbal cues of deception in high-security contexts will be discussed in Chapter Eight.

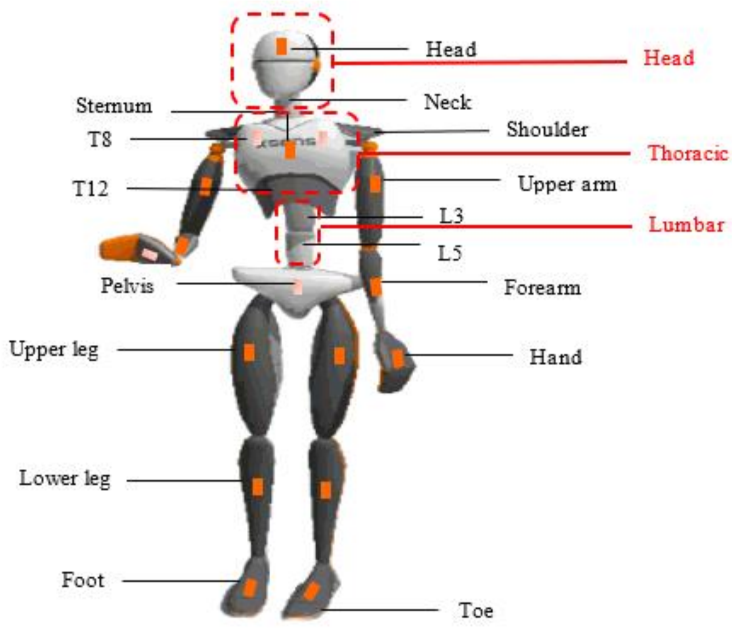
Chapter Three- Methodology

This chapter will detail the four major types of technology featured in this thesis. This includes motion capture and virtual reality (used to different extents in Chapters Four to Seven), point-light displays (Chapter Six) and eye-tracking (Chapter Seven).

3.1 Motion capture

The Xsens™ inertial motion capture system was used to record the non-verbal behaviour which forms the crux of this thesis. Xsens is a wireless system that can capture full-body non-verbal measures. It was chosen primarily due to its precision and precedence in research on non-verbal behaviour (Al-Amri et al., 2018; Ferrari et al., 2010; Van Der Zee et al., 2015). The Xsens suit used in this thesis consisted of a lycra short-sleeved t-shirt, fingerless gloves, footpads, self-fastening straps worn around each limb and the waist, and a set of 17 sensors.

Figure 3.1 Xsens model depicting the sensor placements and the DeLeva segments for analysis



Note. Areas within the red dotted lines of Figure 3.1 are the combined segments (de Leva, 1996) comprised of adjoining joints. The orange rectangles reflect the sensor placement.

Each of the 17 sensors were attached to the participants using either the hook and loop patches on the t-shirt, the self-fastening straps worn around the limbs, the fingerless gloves or the headband. The Xsens system wirelessly records the relative 3D position of 23 joints using the 17 sensors at a frame rate of 120Hz. The sensors can measure the movements of the 23 joints because they are equipped with 3D accelerometers, gyroscopes, magnetometers and barometers which can track: the changes in the position of the body area to which they are attached; when the feet make contact with the ground; the centre of mass position etc. The Xsens MVN™ inertial motion capture system is unable to ascertain absolute positioning. Instead, the sensors measure positioning relative to a global reference frame (the right-handed cartesian coordinate system), wherein the x-axis points to the north, the y-axis points to the west, and the z-axis points upwards (Wu & Cavanagh, 1995).

3.1.1 Recording behaviour using the Xsens system

Each participant wore the Xsens suit with the sensors placed on their: head (attached to the headband); shoulders (attached to the back of the lycra top); sternum (attached to the front of the lycra top); upper arms and forearms (attached to the self-fastening straps); hands (attached to the gloves); pelvis (attached to the self-fastening waistband); upper legs and lower legs (attached to the self-fastening straps); and feet (attached to the footpads).

MVN Animate/Analyze™ software was used in tandem with the Xsens suit and was connected wirelessly. Before recording the non-verbal behaviour, each participant took part in a calibration process to align the position of the sensors to their body and to align the body segments to the global reference frame, so that the measurements were accurate. All participants started the calibration process at the same location within the laboratory, which was denoted as the origin, (0,0,0) in the x,y,z plane. The calibration process involved first assuming a ‘T-pose’, where participants were stationary with their arms horizontally outstretched, followed by walking forwards a short distance, turning 180°, and then walking back to the origin location. Following this calibration process, the non-verbal behaviour could be recorded.

3.1.2 Processing the non-verbal data

The non-verbal behaviour that Xsens measures was exported in the mvnx format, which can be imported into software such as Excel and MATLAB for data processing. Each data file consists of the (x,y,z) coordinates for each of the 23 joints measured, as well as other measures, such as the centre of mass position and timestamps. The exported dataset has n number of rows, dependent on the duration of the recording, and each row represents one frame of the recording.

3.1.3 Analysing the non-verbal data

The automated measurement and analysis of body motion method (AMAB) was used for part of the data processing (Poppe et al., 2014). AMAB is a method of standardising the processing of motion capture data, using MATLAB. The AMAB steps that were used for the non-verbal data in this thesis include the application of a median filter, normalisation in space and normalisation for different participants. These steps will be briefly discussed in the following sections.

3.1.3.1 *Median filter*

As with most technology, the data recorded by the Xsens motion capture suits can be subject to distortions. Magnetic interference and tracking failures can be a source of data distortions with this equipment (Xsens, 2015) and were the cause of some of the data loss reported in Chapter Four (section 4.2.1). However, the occurrence of distortions can be somewhat easier to control in laboratories compared to other settings with certain precautions, such as removing magnetic items in the testing vicinity. The MVN software can smooth distortions in the data by downsampling, which reduces the number of frames within the recording. Using a moving median filter is a technique that can also smooth out distortions. Moving median filters work by systematically working through each data point gained from the Xsens suit and replacing each data point within a *window* (number of frames) with the median of the data points around it. The AMAB protocol (Poppe et al., 2014) works by focusing on the data points within a moving window, between 0.25 to 0.5 seconds, and replacing the values within the window with the median of the values of the neighbouring data points. Using a median filter of this size, as per the suggestion of the authors of the protocol, was deemed appropriate for the analysis of non-verbal behaviour in this thesis because it allows for a happy medium between removing inconsistencies in the data without losing crucial information. Median filtering also reduces the chances of the data smoothing process being skewed by extreme values as can be the case with mean filtering, which can lead to drastic approximations and a loss of important information (Fisher et al., 2003).

3.1.3.2 *Normalisation*

Normalisation is a necessary step to allow for the comparison of non-verbal behaviour between participants. It accounts for variations between participants in their global positioning within the testing space and due to differences in their body size. Two types of normalisation were used: normalisation in space and normalisation for different participants.

3.1.3.2.1 Normalisation in space

The global position of a participant within the laboratory testing space affects the local position of their joints and limbs. To compare movements between participants who may be in slightly different positions in the testing space due to differences in step length etc the joint positions of the participants must be normalised, relative to their own bodies. AMAB sets the pelvis as the 'root' joint, assigning it the location (0,0,0). All other joint and segment positions were calculated relative to the location of the pelvis. For example, the location of the shoulder was calculated by subtracting its global position from the root joint, as opposed to just using the global position of the shoulder.

3.1.3.2.2 Normalisation for different participants

People have bodies of all shapes and sizes. Different limb lengths ultimately lead to differences in non-verbal measures due to their body composition and not necessarily as a result of the experimental manipulation. To make comparisons between participants, the AMAB protocol can scale the limbs, resulting in joint and segment positions that can be compared between participants.

3.1.3.3 *Analysis of non-verbal measures*

Chapter Four assesses the following aspects of non-verbal behaviour: segment displacement, centre of mass displacement, speed, cadence and step length. The reason for the inclusion of these non-verbal measures will be expounded on in Chapter Four. In addition to AMAB, custom scripts were used to analyse these measures, using MATLAB. The following provides a brief explanation of the definition of each measure and how they were calculated.

3.1.3.3.1 Segment displacement

Segment displacement refers to changes in the joint position, relative to the pelvis. This was calculated for each joint j using the Pythagorean theorem. For example, each participants' dataset consisted of a matrix with $f^{\text{rows}} \times 69^{\text{columns}}$ (23 sets of x,y,z coordinates for the joints). Changes in joint position throughout the experiment, between consecutive frames f (rows), were calculated starting with the first frame and ending at the last l, resulting in $f \times 23$ matrices. This was summed and normalised for time, forming an overall amount of movement for each joint. The following equation was used to calculate 3D joint displacement:

$$\delta j_f = \sum_{f=1}^{l} \sqrt{(x_{jf} - x_{-jf+1})^2 + (y_{jf} - y_{jf+1})^2 + (z_{jf} - z_{jf+1})^2} \qquad (1)$$

3.1.3.3.2 Centre of mass displacement

Xsens outputs the positional data of the centre of mass for each frame. Centre of mass displacement was measured throughout the experiment. The change in the centre of mass was calculated as seen in (1), relative to the position of the pelvis which was set to (0,0,0) and the centre of mass positional data that was output by the Xsens system.

3.1.3.3.3 Cadence

Xsens records when the feet make contact with the ground. In the Chapter Four analysis, a step was counted as when either the left foot was flat on the ground whilst the right foot was in the air, or the right foot was flat on the ground whilst the left foot was in the air. The total number of steps per minute was noted as the cadence.

$$Cadence = \frac{Number\ of\ steps}{Time\ (minutes)} \qquad (2)$$

3.1.3.3.4 Speed

Speed varies slightly from velocity because it does not infer direction. Speed was calculated as the change (distance) in the relative position of the pelvis' compared to the origin, using (1), relative to the amount of time that elapsed.

$$Speed = \frac{Distance\ (metres)}{Time\ (seconds)} \qquad (3)$$

3.1.3.3.5 Step length

The distance (length) between the heel position of two consecutive steps reflects the step length. This was calculated using the positional data of each heel strike and averaging the step length for each participant throughout the experiment.

$$Step\ length = \frac{Distance\ (metres)}{Number\ of\ steps} \quad (4)$$

3.2 Virtual reality

Virtual reality can allow greater flexibility in the testing paradigms that can be used. It is an easy way to mimic environments such as security checkpoints which may be difficult to gain access to for field studies, and one can maintain experimental control by simulating environments where extraneous variables may be difficult to control (Brookes et al., 2019).

As discussed in Chapter Two, virtual reality may be used in behavioural research as a method of providing contextual priming/relevance to traditional laboratory studies (de Groot et al., 2020; Parsons, 2015), and may be especially helpful when field experiments are not possible. A common query regarding the use of virtual reality within the behavioural sciences is whether behaviour in virtual environments reflects real-world actions. Nilsson et al., (2018) state that "VR[virtual reality] supports a sensorimotor loop similar to that of the real world and thereby enables users to perceive and act as they would in reality". This statement suggests that functions outside of our conscious control (the sensorimotor loop) determine a level of equivalence between behaviours in and out of virtual reality. The extent to which non-verbal behaviour in virtual reality is reflective of real-life is dependent on two main factors (Nilsson et al., 2018): the ability to move unrestricted and appropriate multi-sensory input. There are three main methods of allowing people to walk whilst in virtual reality. First, by using repositioning systems such as omnidirectional treadmills, second by using proxy gestures such as walking in place and third, featured in this thesis, *redirected walking* (Razzaque et al., 2002; Suma et al., 2015). Redirected walking is a technique used in virtual reality which allows you to physically walk in a virtual environment whilst being unaware of how your movement corresponds to the real world. As a result of this technique, people can travel distances in virtual reality far greater than the small confines of a standard laboratory by walking over the same routes in the real world, which correspond to new areas in the virtual world. Redirected walking has benefits over the other methods of walking in virtual reality because it allows the walker to receive an amount of vestibular information that reflects walking in the real world. For example, walking on a treadmill or in place has been found to reduce the amount of "vestibular self-motion information" compared to redirected walking (Nilsson et al., 2016, 2018). This may be due in part to the lack of corresponding optic flow as well as the different amount of multisensory responses (kinaesthetic, proprioceptive etc) when not taking actual steps due to walking in place, and the additional challenges to postural balance when using omnidirectional treadmills (Chan et al., 2019; LaViola Jr. et al., 2017).

Since this thesis aims to study deception detection relevant to security checkpoints, virtual reality was used to garner responses associated with that environment within the confines of a standard laboratory. There have been contrasting results regarding the similarity in the biomechanics of non-verbal behaviour in virtual environments compared to the real world. Janeh et al. (2017) assessed people walking barefoot in a straight line and found that there were some differences in non-verbal behaviour in virtual reality compared to the real world. However, some of the factors which may have caused these discrepancies including the narrow field of vision (60°) of the head-mounted display (HMD) used and the visuo-proprioceptive information available as a result are easily remedied. Advances in virtual reality HMDs since then, including the larger field of vision, may reduce the occurrence of differences in behaviour. In contrast, Canessa et al., (2019) found that there were no significant differences when people walked in the real world and a virtual environment. Whilst studies have compared gait alone, no studies seem to have compared non-verbal deceptive behaviour in virtual and real environments. However, Bhagavathula et al., (2018) provide a more general insight into the comparability of behaviour in virtual environments and the real world in terms of pedestrian behaviour. Specifically, they tested whether peoples' intentions to cross a road, perception of risk and judgement of car speed differed in a virtual environment compared to a real road. Bhagavathula et al., (2018) found that none of the factors differed between the two environments, except for the perception of the speed of oncoming vehicles.

Although there may be small differences in non-verbal behaviour between virtual reality and the real world, the overall behavioural patterns are similar enough that virtual reality may be a useful tool in tangent with measuring non-verbal behaviour (Fink et al., 2007; Olivier et al., 2018). Specifically, whilst the numerical observations of non-verbal behaviour analysis in virtual reality may need to be approached with caution when generalising beyond the virtual environment, the overall "qualitative behavioral patterns are sufficiently similar to permit tests of theoretical questions in virtual environments" (Fink et al., 2007, p.15). In reference to this thesis, it is advised that the overall presence or lack of a difference between experimental groups take precedence over the quantitative differences in group means, which should be approached with caution when relating the findings to the real world. Concurring with Fink et al., (2007), Soczawa-Stronczyk and Bocian (2020) propose that quantitatively, the differences in virtual and real-world non-verbal behaviour are quite small and may be negligible if using HMDs with a wide field of vision which allow for greater immersion. Furthermore, virtual reality can be a useful tool for assessing biomechanical variation if the isometric mapping of head movements and the first-person virtual camera are incorporated (Nilsson et al., 2018), as was the case in this thesis.

3.2.1 Virtual reality methodology in this thesis

Unity 3D™ was used to create the virtual environments in this thesis. Two types of HMD were used: the HP mixed-reality HMD (Chapter Four); and the Oculus Rift CV1 (Chapters Five and Six). The HP HMD has the distinction of an integrated, internal tracking system. The internal tracking in the HP HMD works due to the accelerometer, gyro sensor, magnetometer and proximity sensor within, which detect the position of the HMD within the virtual and real-world, allowing movements in the real world to align with movements whilst in virtual reality. To make use of the HP HMD's internal tracking, prior to running the virtual program in Chapter Four, a 3×4 metre area within the laboratory was manually tracked with the HMD. The tracking process ensured that the participants' movements within the 3×4 metre space were synchronised with their movements and perception of the virtual environment, as long as they remained within the tracked area. As such, the virtual environment used in Chapter Four was designed to fit within the 3×4 metre space. The Oculus Rift uses external sensors for positional tracking which must be attached to the computer running the virtual program. Compared to the Oculus Rift, the HP HMD allows participants to explore a larger space, beyond the small confines of the area immediately around the computer running the virtual program. Due to this greater freedom of movement, the HP HMD was used in Chapter Four where the participants were required to walk around whilst in the virtual environment because this was not possible with the Oculus Rift. Chapters Five and Six did not require the participants to walk whilst in virtual reality, so the fixed external tracking sensors required by the Oculus Rift was not an issue for these experiments. The Oculus Rift's built-in headphones provide 3D localised audio input which can be beneficial for immersion in the virtual environment (Pollard et al., 2020; Slater et al., 1994). Thus, in Chapters Five and Six the Oculus Rift HMD was used instead.

Much emphasis has been placed on the need for virtual environments to be as immersive as possible to gain results that are reflective of the real world (Slater et al., 1994). Some of the factors which may influence levels of immersion include the visual, auditory and tactile stimuli within the virtual environment. There are limitations in the realism of the visual components of virtual environments. Factors such as the 'uncanny valley' phenomenon (Stein & Ohler, 2017), where avatars that are too realistic can evoke an unsettling feeling, and contrastingly, poor graphics, could also potentially lessen immersion. However, these factors can be difficult to avoid in some situations since the visual stimuli that behavioural researchers use is typically only as good as what is available from technology developers.

Auditory stimuli are also an important aspect of immersion which adds to the ambience of the environment and can contribute to the sense of realism (Slater et al., 1994). Auditory stimuli in the form of background chatter from other avatars and airport audio, gained from recordings of actual airports, were incorporated into the virtual environments in this thesis. Likewise, tactile stimuli can have the same effect. Tactile stimuli were incorporated into the testing environment in Chapter Four, such as when the bag scanner in the virtual environment coincided with a surface in the laboratory, allowing participants to place their bag on something real. This was mentioned by some respondents, in the post-experiment questionnaire in Chapter Four, as having contributed to immersing them in the virtual environment. Figure 3.2 shows a virtual environment similar to that used in the study in Chapter Four, where participants assumed the role of a passenger, whilst Figure 3.3 shows a similar representation of the virtual environment as viewed in Chapters Five to Seven, where participants assumed the role of security personnel and viewed a range of approaching avatars. The C# programming language and Microsoft Visual Studio were used for all the coding relating to the virtual environments in this thesis.

Figure 3.2 Virtual environment from the point of view of a participant in the role of a passenger as in Chapter Four.

Figure 3.3 Sample of a virtual environment from the point of view of a participant in the role of security personnel, as in Chapters Five to Seven.

Note. The avatars in Figure 3.3 are in a t-pose (due to limited access to other images). These postures are not reflective of their movements when viewed by the participants in Chapters Five to Seven.

3.2.1.1 *Avatars*

The humanoid figures that are used in virtual reality are referred to as *avatars* in this thesis. A range of avatars were used (see Appendix A for examples), varying in gender, age, race and height, to be reflective of wider society and the different types of people who frequent travel hubs and security checkpoints. These avatars were created using MakeHuman™ software and were then exported to Unity 3D. Though the bodies of each of the avatars were fully rigged, meaning that their full-body skeletons could be animated, the mouths were not animated due to technological limitations. The abilities required to animate facial movements and synchrony between mouth movements and auditory stimuli were beyond the capabilities available with our existing software and expertise. This is a factor that may have impacted immersion in Chapter Four since participants were being questioned by an avatar.

3.2.1.2 *Avatar movement*

The non-verbal behaviour recorded from the Xsens suits in Chapter Four was used to animate the avatars in Chapters Five to Seven. Avatar animation in this instance refers to using the physical movements gained from the precise joint movements of the participants from Chapter Four who were walking whilst wearing an HMD within a virtual environment. These joint movements were then remapped onto the joints of the avatars used in Chapters Five to Seven so that they moved like the participants in Chapter Four. The non-verbal behaviour recordings were exported from the MVN Animate/Analyze software in the fbx format to Unity 3D, where the rigged avatars were animated. Remapping non-verbal data from Xsens suits to virtual characters in this way has been used within academia (Bideau et al., 2010; Chan et al., 2011) and also within the animated film industry, for example, to create Alice in Wonderland (DeMott, 2010).

Three sets of projects were created in each of the chapters that assessed accuracy (Chapters Five, Six and Seven). The projects were identical except for the non-verbal behaviour that was used to animate the avatars. This allowed for a comparison of the dependent variable (accuracy) with the project number as the independent variable to ensure that any differences in accuracy were due to the experimental manipulation and not due to particularly distinct non-verbal behaviour within the testing stimuli. The projects were not significantly different from each other, allowing the three/four datasets in each chapter to be merged, forming the datasets described in Chapters Five to Seven.

3.3 Point-light displays

Point-light displays feature in Chapter Six. These displays represent non-verbal behaviour, with sixteen joints/segments displayed as lights against a black background. The areas displayed as lights include the head, shoulders, sternum, mid-spine, elbows, wrists, pelvis, hip joints, knees and ankles (see Figure 3.4 for an illustration of how the Xsens non-verbal behaviour translates to a point-light display).

Figure 3.4 Comparison of the Xsens representation of non-verbal behaviour and a point-light display.

When in motion, point-light displays reflect human movement very well, allowing for recognition of factors such as emotion, (Brownlow et al., 1997; Dittrich et al., 1996), gender (Kozlowski & Cutting, 1977; Runeson & Frykholm, 1983), deceptive sports moves (Wright et al., 2013), the direction of movement (Kuhlmeier et al., 2010) and hostile intent (Cohen et al., 2008), amongst others. The point-light displays were shown in 2D following the precedent of the abovementioned point-light display studies which have revolved around the 2D format. Furthermore, showing the point-light displays in 3D format was not deemed to be much more beneficial in aiding deception detection.

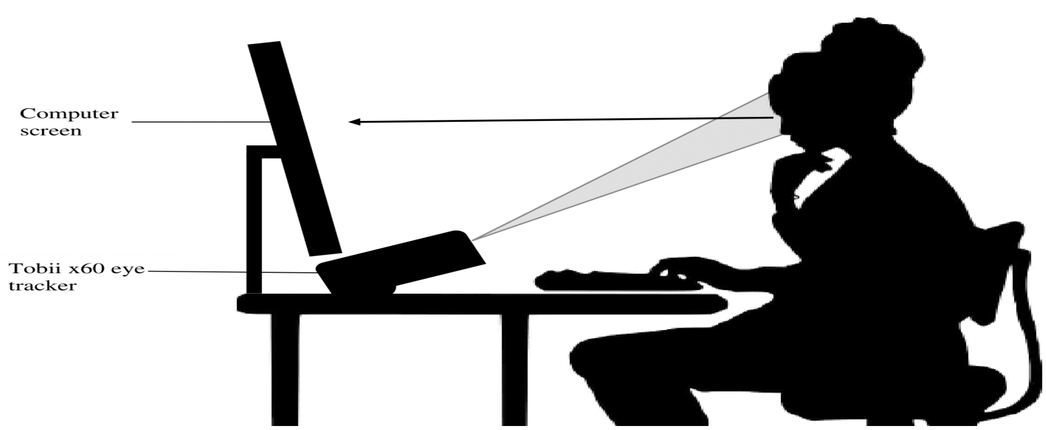
The point-light displays were created using the Xsens recordings gained from Chapter Four and PLAViMoP software (Decatoire et al., 2019). PLAViMoP is open-source software, which interacts with MATLAB and Mokka. This software allows for a standardised method of creating point-light displays. The files containing the non-verbal behaviour data were exported from MVN Animate/Analyze in the c3d format to the PLAViMoP platform. The resultant point-light displays were displayed to participants using the 2D version of Unity, with C# coding used to randomise the video presentation and record participant responses.

3.4 Eye-tracking

The external Tobii x60 eye tracker was used in Chapter Seven. This eye-tracker has a 60Hz sampling rate and was used alongside Tobii Studio software version 3.4.5. This is an unobtrusive eye-tracker, which is placed underneath the computer monitor (see Figure 3.5). Tobii eye-tracking works by using near-infrared light to detect the position of the pupils on the computer screen (TobiiPro, 2015). Eye-trackers such as this are beneficial as they allow people to move their head and eyes in a more natural way (Serchi et al., 2016), compared to more restrained eye-trackers which require a chin-rest. The x60 tracker can detect saccadic eye movements and fixations. Saccadic movements occur when processing visual information, prior to or in between fixations on an object (Bradley, 2014).

To create the eye-tracking stimuli, video recordings were created of the individual avatars walking in a virtual security environment, like those used in the 3D format in Chapters Five and Six. The non-verbal behaviour of the avatars, (honest and deceptive) was gained from the participants in the study in Chapter Four (see section 3.1). Using Camtasia 7 software, the videos were recorded in the AVI format, as required for importation into Tobii Studio. In Tobii Studio the videos were randomised when presented to each participant.

Figure 3.5 Diagram of the eye-tracker setup used in Chapter Seven.



Part II: Chapter Four – Investigating Deception Cues

4.1 Introduction

The field of deception detection is based on the premise that there may be measurable changes in behaviour that can distinguish between honest and deceptive people (Burgoon et al., 2005). Currently, there is a lack of consensus regarding the theoretical grounding of deceptive behaviour, though different approaches of the multi-factor model (Zuckerman et al., 1981) are often cited. Whilst the effects of cognitive load are often taken into account when designing deception studies (Greene et al., 1985; Vrij et al., 2010), it is difficult to replicate the potential "physical and psychological arousal" within laboratories (Stotz et al., 2020, p. 1478) that may be invoked at real security checkpoints. Context could have an impact on the way that deception affects behaviour, therefore researchers should take caution in generalising the results from one setting to another. Few studies have focused on security environments which makes it difficult to determine whether the assumed behavioural differences in deceptive and honest people are relevant to security contexts.

One aspect of the self-presentational theory proposes that honest and deceptive people's behaviour may be equally affected in such contexts. In tangent with the report that some honest Muslim people consciously try to modify their behaviour to ensure that they are perceived well at airports (Blackwood et al., 2015), there may be few differences in non-verbal behaviour, regardless of deceit. If aiming to address an issue that exists within an airport, the context of laboratory experiments needs to try to reflect that as closely as possible, as opposed to arbitrary deception measures which may lack contextual relevance. However, the lack of reliability in the deception cue literature and the Blackwood et al. (2015) findings may hint at the improbability of any individual theory being sufficiently explanatory for all contexts, people and types of deception. The potentially idiosyncratic nature of deception (both between individuals and contexts) may have a detrimental effect on the generalization of lie detection research to the real-world environments in which we hope to identify deception.

Some laboratory research on deception has supported the view that verbal markers of deception are more reliable than non-verbal (Vrij & Fisher, 2020). However, these findings may be impacted by the laboratory-based methodologies/contexts frequently employed when assessing deceptive behaviour. It is becoming apparent that the testing environment needs to take a bigger role in deception research, and specifically when focusing on non-verbal deception, researchers should try to incorporate settings in which non-verbal behaviour may be relied on, such as an airport (Vrij et al., 2019). Vrij and Fisher (2020) suggest that non-verbal measures of deception may be more apparent when using methodologies that are not limited to seated interviews (which is the methodology of choice in a lot of laboratory-based studies). Seated interviews limit the possibility of movement in the deceiver. At security checkpoints, deceivers may engage in a variety of activities and behaviours that are typically restricted when participants are seated and stationary in traditional laboratory-based study designs. Hence, this limitation in the current literature highlights the need to mimic the environments and tasks to which we aim to apply our findings as closely as possible when researching deception. Also, consider that real-world attempts to detect deception/ hostile intent (e.g., a security guard observing someone out of speaking range), may not provide the opportunity for the use of verbal measures, while non-verbal behaviours may be readily observable. Therefore, it is of benefit to investigate whether non-verbal behaviour is distinct when people have criminal intent and consequently whether assessments based on non-verbal behaviour alone are valid.

In recent years, there has been a slight shift towards more studies relevant to airport environments which increase the contextual relevance and applicability of findings to security settings, such as by incorporating replicas of security checkpoints within laboratory settings (Matsumoto et al., 2015) and conducting field studies (Ormerod & Dando, 2014). This study contributes to this movement by using virtual reality to try to closely mimic a security checkpoint environment to potentially induce behaviours relevant to that environment. Using virtual reality allows for more control of the variables than in a field study and can reduce the salience of the laboratory by immersing participants in an experience reflective of a security checkpoint.

Non-verbal behaviour

There are three main aspects of human motor behaviour which can be analysed; kinematic, kinetic and temporal changes (Drillis et al., 1964). Kinematic behaviour can refer to linear displacement measures such as velocity and other positional measures. Kinetics considers the impact of mass/force on the body and includes measures such as centre of mass displacement, whilst temporal measures of body movement refer to changes in body movement relative to time, such as cadence. This chapter focuses on body analysis measures (segment displacement, centre of mass displacement, cadence, speed and step-length) which account for each of these three aspects of human motor behaviour.

One reason why deception may influence non-verbal aspects of behaviour is that people tend to focus more on controlling their facial expressions when deceptive, and as a result, they can be more prone to exhibiting non-verbal cues in other areas of the body while they focus on displaying appropriate facial expressions (Ekman & Friesen, 1969). This reasoning is similar to the behaviour control approach. Aspects of the multi-factor model are often used to explain changes in non-verbal behaviour when people are deceptive, as discussed in Chapter Two. Mullin et al. (2014) found that when people were engaged in deception, they exhibited increases in their postural rigidity which was measured using a force plate. This finding may suggest that deception reduces people's ability to control their non-verbal behaviour, potentially due to the increased cognitive load caused by attempts to maintain a deceptive narrative (Zuckerman et al., 1981). As cognitive load increases, it becomes more difficult to control posture as cognitive resources may be limited due to focusing on maintaining a verbally consistent deceptive narrative (Granhag et al., 2015). Attempts to appear honest may manifest as an overcorrection of behaviour, which is displayed by increased postural rigidity (Mullin et al., 2014). The heightened arousal, caused by deception, may limit people's ability to control their behaviour. This finding spurs the question of whether differences in non-verbal behaviour also manifest throughout walking tasks, in addition to stationary tasks. The non-verbal behaviour in the Mullin et al. (2014) study was assessed solely when standing on a force-plate which may limit the extent to which these non-verbal behaviours would be affected in set-ups where participants took part in a more dynamic task.

In contrast, Matsumoto et al. (2015) is an example of a study that uses a deception task and context similar to the context of interest in this thesis. They assessed deception-related non-verbal changes in a walking task using motion capture technology and built a physical security checkpoint in their laboratory. They found that deceptive individuals displayed differences in non-verbal behaviour, both when on a reconnaissance mission and when carrying a prohibited weapon, compared to a baseline honest walk. For example, velocity and cadence increased at greater rates in the reconnaissance condition compared to the baseline, whereas step length decreased, which may suggest that that deception may limit people's ability to control and maintain regular non-verbal behaviour. These behavioural changes could be interpreted as reflecting a lack of behavioural control wherein normal non-verbal behaviour could not be maintained due to acting deceptively. One reason for the loss of behavioural control is that people are preoccupied with the knowledge that they are in possession of a prohibited weapon/ that they are surveilling the area to ensure the success of their future criminal acts. This preoccupation and the need to conceal their true intention may override their ability to maintain within the normal parameters of non-verbal behaviour. Or akin to the emotional arousal approach of the multi-model theory, deception may have caused an increase in nerves, expressed through their increased cadence and velocity, which could be reflective of trying the complete the deceptive act as soon as possible. Matsumoto et al.'s (2015) research was disseminated as a military report, and unfortunately, their theoretical rationales and conclusions were not expanded on. However, their suggestion that these non-verbal changes may be due to the "psychosomatic influence of deception" (Matsumoto et al., 2015, p.50) aligns particularly with the emotional arousal aspect of the multi-factor theory.

One point of interest that Matsumoto et al.'s (2015) study does not address, is whether different types of deception can be distinguished by non-verbal differences, which a comparison of the weapon and reconnaissance condition could have answered. This is an important area of interest because it would be reasonable to assume that different types of deceit may warrant varying forms of changes in non-verbal behaviour. For example, the salience of physically being in possession of a gun may cause more emotional arousal than being on a reconnaissance mission due to the relative ease of proven wrongdoing if approached when in possession of a gun. Regardless, non-verbal measures may be promising and of value due to instances in the real world where it is the only form of behaviour that can be assessed especially since Matsumoto's findings would suggest that there are effects on the bodily motion of deceivers which can aid us in distinguishing between truth-tellers and liars. Although their use of a physical checkpoint in the laboratory was commendable, it is also possible that this environment still did not sufficiently replicate the contextual pressure associated with being deceptive at a real security checkpoint. The authors note that the "geographic site" (Matsumoto et al., 2015, p.5), amongst other factors, makes broad generalisations difficult. It would be reasonable to assume that the participants were still very aware that they were in a laboratory, responding to questions posed by a research assistant dressed as a security officer. Factors such as the lack of other passengers and relevant audio stimuli etc may have been detrimental to the extent that behavioural responses were reflective of those at a real security checkpoint. To tackle this issue, this chapter uses virtual reality to simulate a security checkpoint. Due to the versatility of virtual reality, using this medium allows for the replication of the context-relevant factors which would be difficult to include otherwise.

Eapen et al. (2010) conducted a motion capture study in which participants were standing, but not walking around. They found that movement reduced when participants lied about their results on a maths test and also when lying about witnessing an accident in the laboratory. Specifically, differences in the overall movement were analysed exclusively in relation to their active verbal deception, but the findings do not reveal whether there were differences in overall movement, simply due to having deceptive intent, regardless of verbal veracity. When assessing people's non-verbal behaviour from afar, as with the SPOT program, security personnel's assessments are reflective of the notion that deception can be identified even when not engaged in an interrogative process. Consequently, though Eapen et al.'s study is insightful, due to the aims of this thesis an investigation of whether there is a continuous impact on non-verbal behaviour is of interest.

The non-verbal measures in this chapter (segment and centre of mass displacement, cadence, step length and speed) were investigated because although they have previously been indicated as distinguishing between honesty and deception, the context of the previous studies and the type of deception hinders the ability to determine whether the same effects would be evident in an airport environment. The non-verbal research was exploratory because the context of studies that used similar measures was too different to this study which used virtual reality and a dynamic task.

Verbal behaviour- Response latency

This chapter will also include a conceptual replication of a response latency study that also used a virtual airport environment (Mapala et al., 2017) and contrasted with other findings of the effect of deception on response latency (Sporer & Schwandt, 2006; Suchotzki et al., 2013). Sheridan and Flowers (2010) found that lying about whether a number was greater than another elicited longer response latencies compared to honest responses, regardless of cognitive load. Conflictingly, Mapala et al. (2017) found that response latencies were shorter when lying about possessing prohibited items in a virtual security environment compared to telling the truth. Regardless of cognitive load, the difference in these results may point toward the importance of the environment and task context in discerning whether lying causes people to have delayed responses, or that people rush to respond, in an attempt to conceal their deceit. The difference in the type of lie, whether one number is greater than another compared to possessing a prohibited item may explain the discrepancy in these findings. The virtual security environment coupled with the subject of the lie may have added pressure and heightened arousal, more so than other methods, resulting in the opposing latency finding. This thesis will assess whether the shorter response latency findings (Mapala et al., 2017) can be replicated. The conceptual replication in this chapter differed from the original study because participants were not seated and there were three deception groups, as opposed to one. As a result of the conflicting findings, the response latency research in this chapter was largely exploratory as the direction of the results was unpredictable.

Response latency was chosen as a dependent variable because it has the potential to be a quick and effective tool for distinguishing between truth and deceit. Other verbal measures of detecting deception, which have better than chance levels of accuracy of correctly identifying deception may be impractical to implement. For example, the controlled cognitive engagement method (CCE; Dando & Ormerod, 2020) relies on open-ended and spontaneous interview procedures, as well as establishing rapport and a baseline with each individual. In airport environments, there is an increasing emphasis on reducing the security screening process for passengers, albeit without compromising safety (Sakano et al., 2016). The relatively low accuracy, 66%, does not justify the amount of time that the CCE warrants, and it still has the potential to fall victim to prejudice-informed assessments since ultimately security personnel are the determiners of people's honesty, whereas response latency could put the onus on people responding within pre-determined parameters. Closed questions were used in this chapter as opposed to open-ended questions because those within the security industry regularly stress that there is a need to reduce the amount of time that it takes for people to be assessed for security, rather than employing more time consuming open interview methods (Gillen & Morrison, 2015; Harris, 2002). Within aviation security, reducing the amount of time that it takes to process through security may also reduce the likelihood that large amounts of people congregate on the airport landside which can become a security target in itself (Martin, 2018; Modić et al., 2018), as was the case in the Brussels and Istanbul attacks in 2016.

This chapter aims to assess whether there are precise and valid measures of deception in a contextually relevant environment. This research has the long-term aim of reducing the likelihood that invalid cues are incorporated into global security detection protocols which are currently used at some security checkpoints, e.g. SPOT, and within interrogations (ACLU, 2017; Denault et al., 2020), by investigating deception in a relevant context. To assess the effect of deception on both non-verbal behaviour and response latency, participants were split into four groups consisting of one honest group and three deceptive groups. The three deceptive groups assessed whether different forms of deception i.e., lying about the possession of prohibited items (weapon or suspicious package) or lying about future intent, would have distinct effects on non-verbal behaviour and response latency. Similar to the weapon focus effect (Loftus et al., 1987) whereby eyewitness testimony may be impacted by the salience of a weapon during the recalled event, lying about having a weapon may result in different changes in behaviour than lying about plans to commit a crime in the near future.

4.2 Method

4.2.1 Participants

An a priori power analysis was conducted using G*Power software (Erdfelder et al., 2009) to determine the sample size that would allow the study to have sufficient power. The significance level was set at the standard $\alpha = .05$, the power level was set to $1-\beta = .80$, and a conservative effect size was set to .25. Accounting for potential data loss due to the use of both virtual reality and motion capture and participant withdrawal, a total sample size of 104 was determined to be sufficient. In total, data was collected from one-hundred and thirty participants (to try to achieve N=104, in light of unusable participant data). Twenty-six percent of the non-verbal data could not be used in the analysis. These exclusions were due to warped data, possibly due to magnetic interference and sensor issues. These rates are not out of the ordinary for this type of research, for example, Eapen et al., (2010) report a 27% data loss rate with their motion capture equipment. Thirty-five percent of the response latency data were excluded from the analysis. These exclusions were due to issues such as unintelligible participant audio, errors with the interviewer audio which may have impacted participant responses, and participants' responses which indicated that they had misunderstood the task.

As a result of the data exclusions, some participants' non-verbal data was usable but not their response latency data and vice-versa. Rather than removing all participants from the sample without both non-verbal and response latency data intact, the non-verbal and response latency data will be reported separately within this chapter. For example, participants whose non-verbal data was not usable due to severe distortions were still included in the response latency sample, because they were unaware of the technology failure, therefore their verbal responses were unaffected. However, participants whose verbal responses indicated that they did not understand the task were excluded from the sample entirely because their movements could not be determined to be representative of their assigned condition.

4.2.1.1 *Non-verbal behaviour*

Ninety-three participants (69% women, 31% male, M_{age} = 20.85 years, SD_{age} = 2.82 years) were included in the non-verbal behaviour analysis. Within this sample, 58% were British and 42% non-British.

4.2.1.2 *Response latency*

Eighty-four participants (66% women, 34% men, M_{age} = 20.68 years, SD_{age} = 2.79 years) were included within the response latency analysis. Within this sample, 62% were British and 37% were 'non-British'.

4.2.2 Sampling

Participants were recruited by responding to posters placed around the Lancaster University campus, or via the university's online recruitment system for psychological studies. All participants were required to be at least eighteen years old, have normal or corrected to normal vision and have no previous adverse reactions to virtual reality. Depending on eligibility, they were either financially compensated or received course credit. Ethical approval was obtained from the Ministry of Defence Research and Ethics Committee and the Lancaster University Faculty of Science and Technology Ethics Committee.

4.2.3 Materials

Data collection took place in a 3×4 metre testing area within a laboratory. An HP mixed reality head-mounted display (HMD) was used as the medium for immersion into the virtual environment. The HP HMD was used as it allowed for positional tracking within the testing area so that participants' movements throughout the 3×4m space were in sync with their perception of their movements in virtual reality. An HDMI extension cord allowed participants to walk around the entirety of the testing area whilst wearing the HMD.

4.2.3.1 *Conditions*

There were four conditions: *honest*, *future action*, *package* and *weapon*. Those in the honest condition were truthful for the duration of the experiment. Those in the future action condition were deceptive about intending to meet with someone once they had passed through the security area to commit a prohibited act, whereas those in the package and weapon condition were deceptive about one of the items in their possession. The conditions were randomly assigned.

4.2.3.2 *Items*

All participants were provided with an empty backpack and a range of items. The items included pens, a phone, a notepad, a magazine and a small case. An additional condition-specific item was provided. The honest and future action conditions were given a fiction book as their additional item to ensure that all participants had the same number of items. Those in the weapon condition were provided with a deactivated handgun, and those in the package condition were given a sealed package filled with a powdery substance. Participants in each condition were verbally made aware of which of their items were prohibited (if any).

4.2.3.3 *Virtual reality*

The virtual environment was created using Unity™, to emulate a high-security environment. This environment included a metal detector, bag scanner, and auditory stimuli to mimic the ambience of a security checkpoint as well as a range of virtual avatars such as animated security guards, background characters and airport staff. The humanoid virtual characters were created using MakeHuman™.

4.2.4 Measures

4.2.4.1 *Non-verbal measures*

For the duration of the experiment, participants' movements were recorded using the Xsens™ suit via the 17 sensors which were placed on the body to measure various aspects of non-verbal behaviour (see Chapter Three). The non-verbal measures were segment displacement, centre of mass displacement, speed, cadence and step length. Non-verbal behaviour was recorded for the duration of the task mainly due to it being impractical for participants to wear the suit solely during the interview portion of the study because of time constraints and this would have impacted their immersion in virtual reality. Furthermore, wearing the suits throughout allowed us to assess whether there were condition related changes in non-verbal behaviour regardless of active vocal deception, which was of interest due to the greater potential for integration with automated methods of detecting deception.

4.2.4.2 *Response latency interview*

Participants were asked a series of 16 questions, adapted from Mullin et al (2014; see Appendix C), in random order, regarding their intentions and the content of their bag. The questions were asked by a female virtual avatar, who spoke in a southern English accent, using text-to-speech software. All participants were asked the same set of questions which required a verbal "yes" or "no" response. Two response types were garnered by the questions, either: an honest response or a genuinely deceptive response about an item or intention. Participants' responses were in accordance with their assigned conditions. Those in the deceptive conditions were instructed to conceal that they had a prohibited item or that they were intending to commit a prohibited act once through the security check. The following examples illustrate how responses differed depending on the condition. "Do you have any prohibited items in your bag?", garnered an honest "no" response from those in both the honest and future action condition, as neither had prohibited items. Yet, this same question garnered a deceptive "no" response from those in the weapon and package condition, to conceal that they had prohibited items in their bag. "Will you be meeting with anyone else?" would have elicited an honest "no" response from those in the honest, package and weapon condition, and a deceptive "no" from those in the future action condition, as they were instructed to meet up with someone once past the security checkpoint. The question “Did you pack your bag yourself?” elicited an honest “yes” response from all four conditions.

Using closed questions allows for potential future automation of this type of measure, which could reduce the impact of bias which may be inherent with open questions delivered by an interviewer, who ultimately has the power to decide whether someone may pose a threat.

4.2.4.2.1 Audio measurement

Participants’ verbal responses to the 16 questions were recorded whilst they were in virtual reality, using OBS Studio™. The audio data was exported to Audacity™ to measure response latency which allows for audio measurements of the latency between the end of each of the interviewer’s questions and the beginning of participants’ responses, to the nearest millisecond.

4.2.4.3 *Post-experiment questionnaire*

The Edinburgh handedness questionnaire (Oldfield, 1971) was initially included to see whether handedness and condition affected how participants carried the bag. However, to avoid interfering with the motion capture sensors, all participants held the bag in the same way, so it was not included in the analysis. An immersion questionnaire (Appendix B), adapted from Usoh et al., (2000), quantified how immersed participants were in the virtual environment, with potential immersion scores ranging from zero to 30. A standard demographics questionnaire was also administered.

4.2.5 Design

This experiment used a mixed measures design, with condition (honest, future action, package and weapon) as the between factor and response type (honest or deceptive) as the within factor. Condition and response type were the independent (predictor) variables. The non-verbal measures (segment displacement, centre of mass displacement, speed, cadence and step length) and response latency were the dependent variables. Note that the response type variable was not included as an independent variable in the non-verbal analysis.

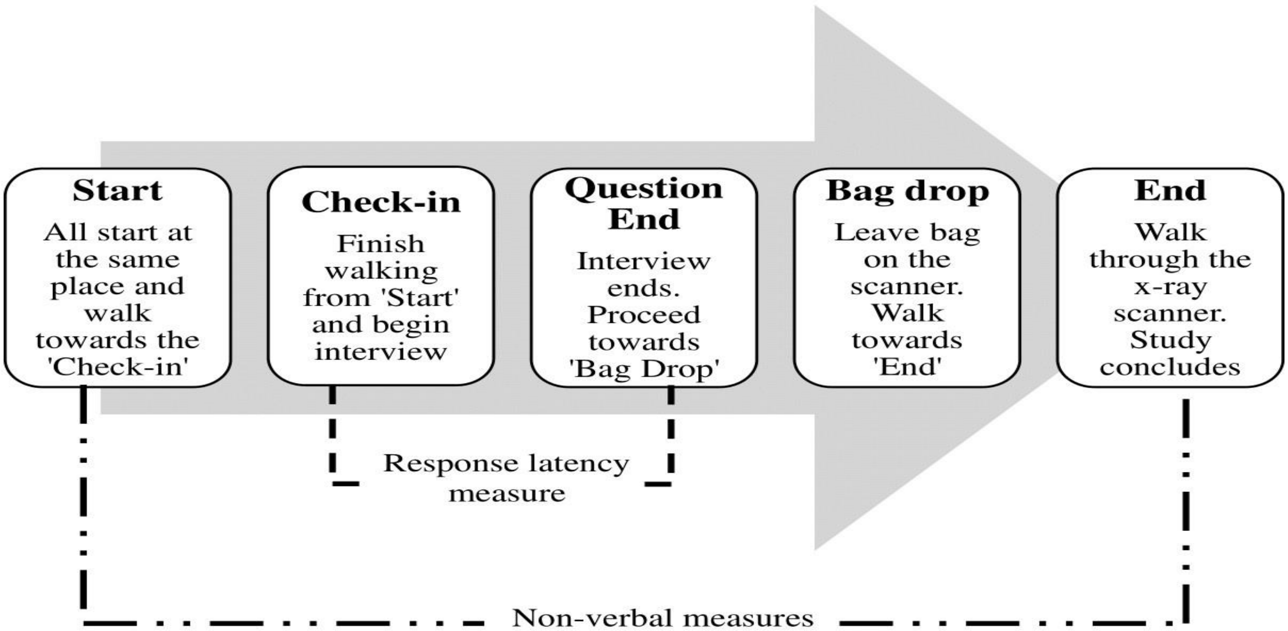
4.2.6 Procedure

Informed written consent was gained before the commencement of the study. Participants wore the Xsens suit, which consisted of a lycra t-shirt and 17 sensors (see Chapter Three, Figure 3.1). Each participant took part in a motion capture calibration process which synchronised the sensors to their body movement. Then, they were given instructions concerning the experimental task, as appropriate for their assigned condition, with their main task being to successfully proceed through a security checkpoint. Participants were told that they would be in a virtual security checkpoint environment, where they needed to first walk up to a desk, then convincingly respond to a series of questions without letting the interviewer know if they were in possession of a prohibited item or that they were being deceptive about their plans. In the honest condition, participants were simply instructed to respond to the interview questions truthfully. Participants were informed that following the series of questions, they would be prompted by the interviewer to walk to the security scanner to place their bag down on the conveyor belt. Then they were to walk through an x-ray scanner and past a security guard to collect their bag, after which they would be through the security checkpoint. Those in the future action condition were informed that once past this point, they would be approached by another avatar with whom they would commit a prohibited act. Participants in the future action condition could see the avatar that they were to meet up with, but the study concluded before they interacted. The future action condition merely required the participants to have the intention of meeting up with someone.

Following the instructions, all participants were provided with a set of items, per their condition, and provided with a bag into which they placed the items, see section 4.2.3.2. The items, notably the deactivated handgun and the unidentifiable suspicious package were explained to be regarded as prohibited items. Likewise, those in the future action condition were briefed that they were to conceal that they were going to meet up with someone (who was visible in the background) with whom they would commit a prohibited act once through security. All participants were given a final verbal prompt to ensure that they had understood their task and encourage any last-minute questions. Participants then picked up the packed bag, and the HMD was placed on their head, immersing them into the virtual environment. All were provided with a set of headphones to use with the HP-mixed reality HMD which does not have built-in headphones.

Once the participants were in the virtual environment, they walked to a check-in desk where they answered a series of 16 randomised interview questions. The questions were randomised to ensure that any differences in response latency were not due to the order that the questions were asked. Following the interview, they walked to a bag scanner on which they placed their bag. Then they walked through the x-ray scanner, which concluded the study. The Xsens suit was removed, they completed the post-study questionnaires and were debriefed.

Figure 4.1 Five key stages of the experimental procedure whilst in virtual reality, during which the non-verbal behaviour and response latency were measured



4.2.7 Analysis

The analyses were conducted using R Studio (RStudio, 2020).

4.2.7.1 *Non-verbal behaviour*

The non-verbal data was first processed using the AMAB system (Poppe et al., 2014) and then using custom scripts developed in MATLAB. . Condition (honest, future action, package and weapon) was the independent variable. Two MANOVAs were used. In the first MANOVA, there were 15 segment displacement dependent variables (see the DeLeva segments, Chapter Three), while the second MANOVA had five dependent variables (overall movement per minute, cadence, step length, speed and centre of mass displacement). Within both MANOVAs were three stages of analysis which assessed: 1) the differences between the honest condition and the three deceptive conditions. 2) the differences between the future action condition compared to the package and weapon conditions, and 3) the package condition was compared to the weapon condition.

4.2.7.2 *Response latency*

The lmerTest package (Kuznetsova et al., 2017) was used to produce a linear mixed model with *p*-values. Response latencies were not log-transformed, to allow for direct interpretation of the results, which may have been hindered by log transformations (Lo and Andrews, 2015). Condition (honest, future action, package and weapon) and response type (honest or deceptive) were the predictor (independent) variables. Response latency was the dependent variable.

4.3 Results

4.3.1 Non-verbal results

The first MANOVA analysed displacement between the 15 DeLeva segments. Table 4.1 shows that there were no significant differences in segment displacement between the honest and deceptive conditions.

Table 4.1 Segment displacement MANOVA comparing the honest condition to the three deceptive conditions

Segment	$F(3,91)$	η_p^2	p
Head	0.71	.02	.55
Lumbar	0.36	.01	.78
Thoracic	0.60	.02	.61
Right upper arm	0.75	.02	.53
Right forearm	0.75	.02	.52
Right hand	0.71	.02	.55
Left upper arm	1.26	.04	.29
Left forearm	1.01	.03	.39
Left hand	1.16	.04	.33
Right upper leg	0.65	.02	.59
Right lower leg	0.92	.03	.44

Right foot	1.27	.04	.29
Left upper leg	0.74	.02	.53
Left lower leg	0.68	.02	.57
Left foot	0.85	.03	.47

Table 4.2 shows that were no significant differences in segment displacement in the future action condition compared to the package and weapon conditions.

Table 4.2 Segment displacement MANOVA comparing the future action condition to the package and weapon conditions

Segment	$F(2,67)$	η_p^2	p
Head	1.26	.04	.29
Lumbar	0.50	.01	.61
Thoracic	1.03	.03	.36
Right upper arm	1.23	.04	.30
Right forearm	1.15	.03	.32
Right hand	1.07	.03	.35
Left upper arm	2.04	.06	.14
Left forearm	1.88	.05	.16
Left hand	2.34	.07	.10
Right upper leg	.99	.03	.38
Right lower leg	1.53	.04	.22
Right foot	2.04	.06	.14
Left upper leg	1.51	.04	.22
Left lower leg	1.20	.03	.31
Left foot	1.30	.04	.28

Table 4.3 shows that there were some significant differences in segment displacement between the package and weapon condition, notably in the left forearm and left hand. The other segments were not significant.

Table 4.3 Segment displacement MANOVA comparing the package and weapon condition

Segment	$F(1,46)$	η_p^2	p
Head	1.85	.04	.18
Lumbar	0.69	.01	.41
Thoracic	1.54	.03	.22
Right upper arm	2.23	.05	.14
Right forearm	2.22	.05	.14
Right hand	2.48	.05	.12
Left upper arm	2.84	.06	.10
Left forearm	4.27	.08	.04
Left hand	4.90	.10	.03
Right upper leg	0.35	.007	.56
Right lower leg	1.08	.02	.30
Right foot	1.37	.03	.25
Left upper leg	0.71	.02	.40
Left lower leg	1.08	.02	.31
Left foot	0.53	.01	.47

Table 4.4 Means and standard deviations of the segment displacement between conditions

Segment	Condition							
	Honest		Future action		Package		Weapon	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Head	8.25	1.72	8.17	1.16	8.16	1.40	8.67	1.20
Lumbar	2.20	0.54	2.20	0.38	2.21	0.38	2.30	0.36
Thoracic	11.70	2.41	11.70	1.71	11.60	2.00	12.30	1.75
Right upper arm	3.67	0.76	3.65	0.67	3.36	0.58	3.88	0.58
Right forearm	2.98	0.85	2.93	0.95	2.96	0.45	3.23	0.75
Right hand	3.84	1.33	3.80	1.52	3.74	0.55	4.21	1.36
Left upper arm	3.43	0.72	3.39	0.54	3.40	0.53	3.69	0.66
Left forearm	2.83	0.95	2.70	0.82	2.64*	0.35	2.99*	0.74
Left hand	4.13	1.63	3.80	1.34	3.85*	0.61	4.38*	0.99
Right upper leg	0.91	0.28	0.87	0.25	0.93	0.25	0.97	0.24
Right lower leg	3.46	0.72	3.32	0.54	3.44	0.75	3.63	0.49
Right foot	9.90	2.13	9.19	1.85	9.65	1.83	10.20	1.52
Left upper leg	0.96	0.47	0.85	0.23	0.91	0.22	0.97	0.28
Left lower leg	4.55	1.44	4.22	0.64	4.32	0.84	4.54	0.62
Left foot	17.1	3.16	16.1	2.29	16.8	3.48	17.4	2.16

Note. The asterisks indicate the conditions that were significantly different in the displacement of a segment.

The second MANOVA revealed no significant differences between the honest and deceptive conditions, in overall movement per minute $F(3,89) = 1.38$, $p = .25$, $\eta_p^2 = .04$ cadence $F(3,89) = 1.00$, $p = .40$, $\eta_p^2 = .03$, step length $F(3,89) = 0.93$, $p = .43$, $\eta_p^2 = .03$, speed $F(3,89) = 1.03$, $p = .38$, $\eta_p^2 = .03$, or centre of mass displacement $F(3,89) = 0.46$, $p = .71$, $\eta_p^2 = .02$.

There were no significant differences in the future action condition compared to the package and weapon conditions in overall movement per minute $F(2,65) = 2.45$, $p = .09$, $\eta_p^2 = .07$ cadence $F(2,65) = 0.99$, $p = .38$, $\eta_p^2 = .03$, step length $F(2,65) = 1.25$, $p = .29$, $\eta_p^2 = .04$, speed $F(2,65) = 1.47$, $p = .23$, $\eta_p^2 = .04$ or centre of mass displacement $F(2,65) = 0.35$, $p = .70$, $\eta_p^2 =.01$.

There were no significant differences between the package and weapon conditions. Overall movement per minute $F(1,46) = 2.20$, $p = .15$, $\eta_p^2 = .05$ cadence $F(1,46) = 1.48$, $p = .23$, $\eta_p^2 = .03$, step length $F(1,46) = 2.02$, $p = .16$, $\eta_p^2 = .04$, speed $F(1,46) = 2.12$, $p = .15$, $\eta_p^2 = .04$, or centre of mass displacement $F(1,46) = 0.18$, $p = .68$, $\eta_p^2 = .004$. Immersion scores for the non-verbal subsection of participants was good, ($M = 20.69$, $SD = 5.18$), at a rate of 68%.

Table 4.5 Means and standard deviations of movement per minute, cadence, step length, speed and centre of mass displacement

Measure	Condition							
	Honest		Future action		Package		Weapon	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Overall movement per minute	80.00	16.00	76.9	13.00	78.3	12.30	83.4	11.5
Cadence	7.21	2.81	7.11	2.62	7.11	2.62	6.21	2.52
Step length	0.50	0.28	0.52	0.20	0.48	0.23	0.60	0.34
Speed	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
Centre of mass displacement	0.00	0.01	0.00	0.01	0.06	0.60	0.06	0.60

4.3.2 Response latency results

Table 4.6 shows the linear mixed model with response latency as the dependent variable, condition and response type as the predictor variables. The individual questions and participant identification number were included random factors (see Table 4.7).

Table 4.6 Linear mixed model with condition (honest condition as the intercept) and response type (honest response as the reference level) as predictor variables

Predicted variable	Estimate	95% CIs		*t*	*p*
		Lower	Upper		
(Intercept)	769.86	694.36	845.31	19.65	<.001
Future action	1160.43	59.36	261.47	3.08	<.01
Package	41.86	-58.77	142.55	0.81	.42
Weapon	82.21	-14.86	179.24	1.64	.10
Response type (deceptive)	-37.84	-110.30	34.67	-1.02	.32
Future action× Response type (deceptive)	-34.08	-108.18	40.21	-0.90	.37
Package× Response type (deceptive)	21.97	-51.53	95.39	0.59	.56
Weapon× Response type (deceptive)	-25.41	-95.84	45.01	-0.71	.48

Table 4.7 Random effects for the linear mixed model in Table 4.6

Predicted variable	Standard deviation
Participant	151.54
Question	49.37
Residual	212.58

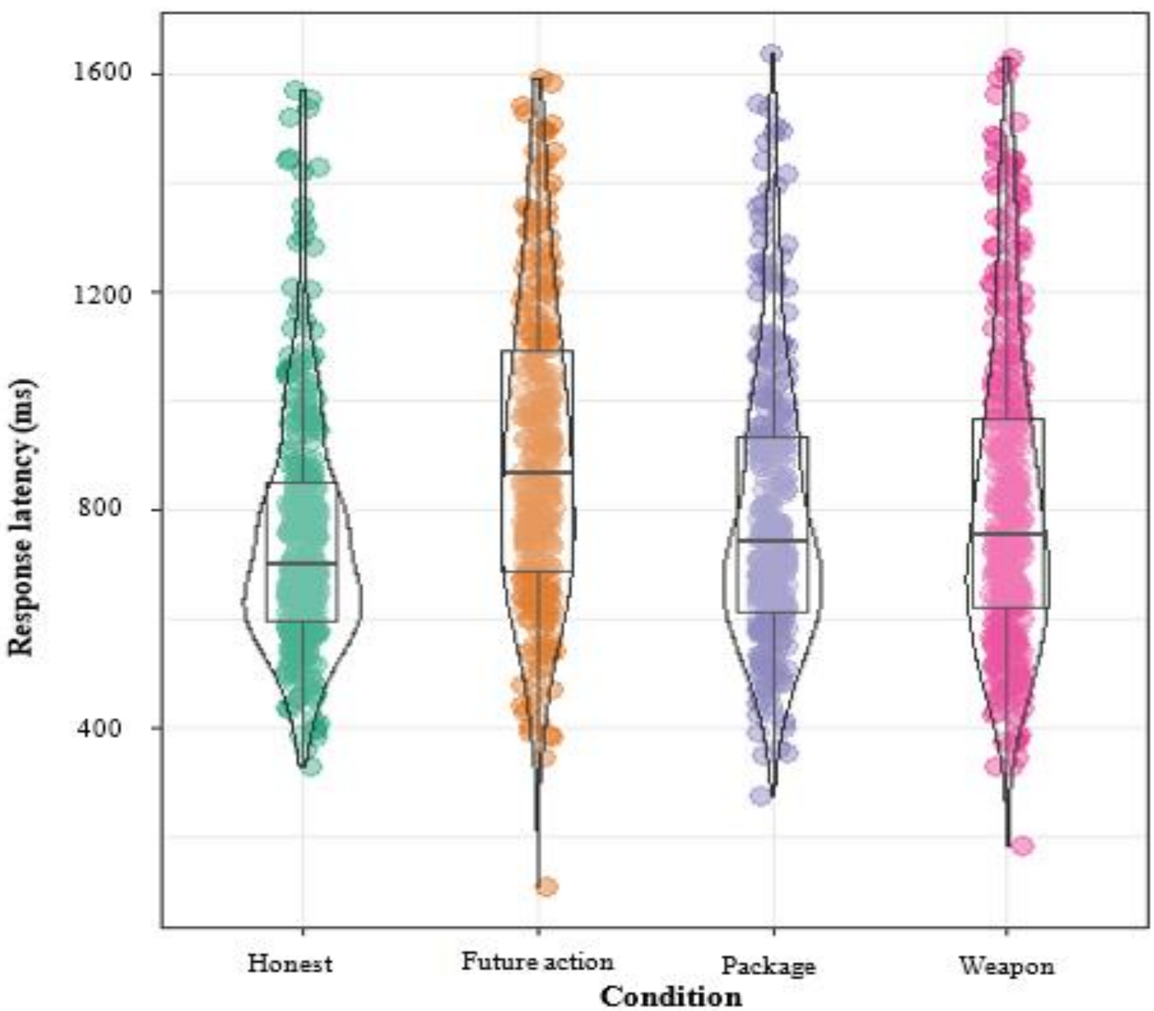
Figure 4.2 shows the difference in response latency between the four different conditions. Honest condition (M = 752.60, SD = 226.22), future action (M = 897.65, SD = 275.62), package (M = 795.31, SD = 256.99) and weapon (M = 816.89, SD = 284.75). There was no significant difference in response latency based on response type; honest response (M = 827.33, SD = 264.27), deceptive response (M = 782.31, SD = 269.40) and there was no interaction between

condition and response type (see Table 4.6 and Table 4.8). Immersion scores for the response latency subsection of participants were good (*M* = 21.43, *SD* = 4.73), a rate of 71. 43%.

Table 4.8 Descriptive summary of the response latencies with the response type

Condition	Honest response		Deceptive response	
	M	*SD*	*M*	*SD*
Honest	764.03	223.29	727.29	231.72
Future action	918.57	275.38	854.39	272.56
Package	800.67	251.70	783.22	269.58
Weapon	837.41	281.32	771.99	288.35

Figure 4.2 Violin plot of the response latencies for the honest, future action, package and weapon condition



4.4 Discussion

4.4.1 Non-verbal behaviour discussion

Research on more holistic assessments of non-verbal behaviour within the deception field has suggested that there are tangible changes in people's movements that indicate deception. However, the context of these studies made it difficult to determine whether similar patterns of behaviour would be seen in a virtual airport setting compared to a seated or standing (stationary) laboratory interrogation setting. The lack of significant differences of this study would suggest that non-verbal behaviour did not distinguish honesty from deception, though there were differences in the left forearm and hand between two of the deceptive conditions, package and weapon.

Assessing segment displacement as a measure of deception detection using motion tracking suits is a relatively new area of research. Although non-verbal measures of deception have been investigated in the past, those using motion capture suits have tended to focus on variables other than segment displacement, such as spectral frequency (Matsumoto et al., 2015) and angular changes (Randhavane et al., 2019). This chapter did not find significant differences in non-verbal behaviour between the honest and deceptive conditions, which may cast doubt on the utility of non-verbal behaviour measures and observational judgements of deception at security checkpoints as used presently, in some places. Although there were differences between two of the deceptive conditions, as will be discussed, the primary need at security checkpoints is for tools that can distinguish between honest people from those who may have deceptive intent and potentially pose a security risk. This lack of a significant finding in segment displacement contrasts with (Van Der Zee et al., 2015) who found that people had more displacement (moved more) when they were lying and (Eapen et al., 2010) who found the opposite effect, that people moved less.

Similarly, the lack of significant differences in overall movement per minute, cadence, step length, speed or centre of mass displacement may suggest that non-verbal behaviour may not be a suitable determiner of intent. These measures, in addition to segment displacement, were selected to provide a well-rounded measure of non-verbal behaviour, given that they encompass kinetic, kinematic and temporal aspects of gait. The null findings conflict with the literature (Matsumoto et al., 2015; Mullin et al., 2014; Yu et al., 2015) and may be attributed to factors such as the lying task itself and the context of the task. For example, the main difference with the Matsumoto task is the use of virtual reality whereas the main difference with Mullin in addition to the virtual airport is that their task did not require participants to walk. Regardless, these three vastly different results show the volatility of the use of non-verbal behaviour to determine intent given the potential for context and the subject of the lie to drastically result in different patterns, or lack thereof, in non-verbal behaviour.

Future studies may wish to consider incorporating a block design that allows for analysis of non-verbal measures per response type which may gain further insight into whether there are any changes in non-verbal behaviour when actively deceptive using a walking task. The random order of the questions in this study makes it difficult to determine this. Also, the technological issues were a hindrance in the data collection process, so other methods of measuring full-body non-verbal behaviour may need to be explored to make data collection more efficient.

4.4.2 Response latency discussion

The response latency results are somewhat in agreement with the majority of the literature (Sporer & Schwandt, 2006; Suchotzki et al., 2013; Walczyk et al., 2003) as there was a significantly shorter response latency in the honest condition compared to the deceptive future action condition, conflicting with previous research which used a similar methodology (Mapala et al., 2017). It is difficult to ascertain why there were no significant differences in response latency with the weapon and package condition. Perhaps, it was easier to lie about being in possession of these items, whereas lying about future intent may have required more thought about what information could be revealed.

There was no within-subject effect of response type, which conflicts with previous findings which suggested that response type affects response latencies within conditions (Mullin et al., 2014), with shorter response latencies to questions when actively lying compared to honest responses (Mapala et al., 2017; Vrij, Fisher, et al., 2008). The lack of a response type effect, conflicting with previous virtual reality research (Mapala et al., 2017), may imply that having deceptive intent is taxing on overall cognitive processes to the extent that there is no distinguishable difference on response latency when actively lying, compared to honest responses. Having deceptive intent may induce a general state of apprehension which is present even when telling the truth, requiring extra time to ensure that an honest response would not reveal any compromising information. It is possible that the finding of shorter response latencies when lying was unique to that study (Mapala et al., 2017). Vrij & Fisher (2020) propose that cues to deception gained in a seated interview scenario, as was the case in the conflicting research (Mapala et al., 2017), may differ from those gained when walking or in other situations, which may explain these findings.

4.4.3 Limitations

An intentional decision was made to use a design whereby participants were assigned to one condition only. This was chosen with the potential applicability in mind. If non-verbal measures were to be used in a real setting, it is unlikely that previous honest baseline measures would be available to compare to their present behaviour. Consequently, there is value in assessing whether the mean non-verbal measures of an honest population differ from a population with deceptive intent. Furthermore, to address the use of observational non-verbal assessment used by some organisations (Denault et al., 2020), which this thesis aims to do, it was necessary to try to replicate their assessment environment. It could be assumed that when security personnel assess someone's non-verbal behaviour for signs of deception, they are comparing them to a perceived norm or checklist (Committee on Science, Space and Technology, 2011) of how honest people behave generally, not how that specific individual behaves normally. There may have been significant differences in non-verbal behaviour in this study when comparing participants' baseline to their behaviour when deceptive, however, that type of study design was not deemed helpful considering the aims of this thesis to mimic procedures in which non-verbal baselines cannot be established on an individual basis.

Although virtual reality was used to narrow the gap between traditional laboratory experiments and the real world, it is possible that virtual environments affected the outcome variables. Immersion scores were good relative to the maximum immersion score possible; however, it should not be understated that this is one of the few studies where participants' non-verbal behaviour was assessed whilst they were walking in virtual reality and the real world. As such it is difficult to judge whether there was an impact of virtual reality on participants' non-verbal behaviour, separate from the effect that deception may have had, as there is little to compare in the absence of a non-virtual condition. Similarly, it is difficult to ascertain the extent to which the virtual environment was able to coerce behaviour that is reflective of behaviour at a real security checkpoint. Although previous research in other domains has found similar behavioural responses in virtual reality and real-life (Bhagavathula et al., 2018), future comparison research may wish to assess non-verbal in a virtual scenario, field experiment and traditional, non-virtual laboratory experiment using an identical procedure, to see whether there is an effect/interaction with the testing paradigm on these non-verbal and verbal measures in these three settings.

The relative space constraints of the laboratory may have impacted the speed with which people walked, contributing to the null results of deception on speed, which contrasted with Matsumoto et al.'s (2015) findings that liars walked faster than those telling the truth. Had the space been larger, and therefore the amount of time that participants were observed and had to walk towards the interviewer was longer, perhaps this would have led to differences in speed between the conditions. Wearing the HMD may also have contributed to their walking speed whilst in virtual reality. People may have been apprehensive of walking fast because they could not see the real environment that they were walking in, though Canessa et al., (2019) did not find similar differences caused by wearing an HMD.

The limited age range of participants may limit the generalisability beyond the relatively young sample used in this study. Much of the literature on the effects of deception points to cognitive load as one of the major causes of changes to non-verbal and verbal behaviour (Vrij, Fisher, et al., 2008; Zuckerman et al., 1981). However, not much focus is placed on "non-pathological cognitive ageing" (Deary et al., 2009, p.137) which can impact our ability to handle the increased cognitive load, with declines in aspects of cognitive functioning, e.g. processing speed and executive functioning starting from the age of 30. Cognitive load can affect gait in increasing amounts as people age (Verrel et al., 2009), so a sample with a wider age range would be of value in future research to see if there are interaction effects with age and deception or if the findings from this chapter are applicable over a wide range of ages.

4.4.4 Practical application

These findings are important as they cast doubt on the use of non-verbal behaviour as a method of identifying those who may pose security threats within some security checkpoint protocols (ACLU, 2017; Weinberger, 2010). These findings should contribute to pushing for more contextually relevant research on the impact of deception on non-verbal and verbal behaviour to help paint a clearer picture of the potential validity of current deception detection tools and the development of new ones. With the aim of working towards establishing valid security protocols, more research relevant to an airport environment could help to ensure that: screening for deceit is a more efficient process; fewer deceptive people slip through the cracks; fewer honest people are unnecessarily subjected to prolonged interrogation (Hashad, 2004).

Xsens motion capture suits were used in this study to establish whether non-verbal behaviour was of importance as a deception cue, as it is an easy system to use in laboratory environments. There are other aspects of non-verbal behaviour which could be assessed in future research, such as joint angle differences, which may show differences between honest and deceptive people. Though not warranted based on the findings of this study, were future studies more fruitful, a non-intrusive system, much like Bouchrika et al's., (2011) CCTV pose matching, or blob analysis (Lu et al., 2005) with the capability to analyse aspects of non-verbal behaviour in real-time may be necessary for investigation and implementation in the high-volume areas where security checkpoints are located. Similarly, any potential use of response latency measures would also require the development of a quick ‘hands-off’ system for measuring latencies, should future studies find differences between other types of deception.

The response latency linear model shows that there are only differences between the honest and future action conditions. Considering these findings, response latency would not be a good measure to introduce at security checkpoints given that other forms of deception are not distinguishable from honesty using this measure. This could potentially result in false negatives and/or false positives in an environment where precision is very important due to the impact it can have on both efficiency and the effect on those being assessed.

4.4.5 Summary

This chapter has not found significant differences in non-verbal behaviour between honest and deceptive people when using virtual reality to simulate an airport environment and automated measurement tools. Consequently, the findings in this chapter do not support the use of these non-verbal measures or response latency to distinguish between honest and deceptive people.

Part III: Chapter Five – Deception Identification Accuracy and Decision-Making

5.1 Introduction

The null findings of the non-verbal measures in Chapter Four could be reflective of similar behavioural patterns in honest and deceptive people at security checkpoints. If so, then how accurate are people at identifying deception when the non-verbal measures from the previous chapter are not significantly different and given the unreliability of non-verbal markers generally, as suggested by the literature (DePaulo et al., 2003; Luke, 2019)? Also, in these instances, what factors are relied on to inform decision-making and ultimately how valid is the inclusion of non-verbal measures in security protocols? (ACLU, 2017; Denault, 2020; Jupe & Denault, 2019; U. S. Government Accountability Office, 2017).

Non-verbal assessments of deception at real-world security checkpoints often rely on naked-eye observation of others' behaviour. By default, the inclusion of non-verbal measures in security protocols such as SPOT, is contingent on the presence of consistent and observable differences in non-verbal behaviour, to identify those who may pose a threat. Chapter Four used virtual reality as a contextual primer with the notion that participants may exhibit behaviours like they would at an actual airport. Additionally, virtual reality is used in this chapter to help to provide an assessment of the validity of using non-verbal behavioural measures in airport settings.

Whilst it is important to research if there are valid cues that indicate deception, as in Chapter Four, it is equally important to gain an understanding of people's ability to identify deception and the factors that they consider when judging if others are deceptive. Brunswik's lens model (Brunswik, 1952) illustrates the need for research to approach deception from two viewpoints: that of the deceiver and also the person observing the deception. Consequently, the study presented in this chapter will focus on the decision-making behaviour of an observer of deception and their accuracy at identifying honesty and deception.

It is important to note that although the previous chapter focused on some non-verbal behaviour related differences, there are other non-verbal aspects of behaviour that may be relied on to inform observational judgements of deception, as mentioned in Chapter Two, more than can be investigated in one thesis. Consequently, although the non-verbal measures being assessed were not significant, there may be other non-verbal factors that people use to inform their decisions. People who work within the frontlines of the security industry sometimes remark about having an intuition, or "sixth sense" (Pinizzotto et al., 2004, p.6) that aids them in distinguishing honesty from deceit.

Proponents of the inclusion of non-verbal behaviour markers in security programs suggest that security personnel can identify deceit (Davis et al., 2013; Ekman, 2011), despite the lack of consistent non-verbal markers and dubious scientific justification of their inclusion (U. S. Government. Accountability Office, 2010). Accuracy of detecting deception has been studied a lot, with meta-analyses revealing an overall 54% accuracy rate (Bond & DePaulo, 2006), with an average underlying distribution of 47% correct identification of deceit and 61% identification of honesty. The potential for deception identification errors may have severe ramifications in certain contexts, as evidenced within police interrogations (Baldwin, 1993) where falsely presumed guilt or innocence can spiral from an interrogation into long and undeserved prison sentences wherein, the true culprit goes free (Benner, 2009; InnocenceProject, 2019). Likewise, at security checkpoints, the accuracy of detecting deceit may be impacted by preconceived notions of the profile of someone wishing to commit an illegal act at such a location (Press, 2010). A "loss of individualization" of deceptive people (Harcourt, 2006, p. 109), reflected by applying intense scrutiny at airports to specific populations, can hinder deception detection rates by producing an inflated belief that the remaining population is unlikely to commit criminal acts (Press, 2010).

While a lot of the literature in this field has focused on accuracy concerning assessments of verbal statements (Frank et al., 2008; Hartwig et al., 2006; Vrij et al., 2011), few have focused purely on veracity assessments based on non-verbal measures, yet often judgements in airport contexts are initially based on non-verbal behaviour. Runeson & Frykholm (1983) lay the groundwork for the premise that deception can be ascertained from non-verbal behaviour alone. They found that observers were able to identify when people were faking the weight of a box they were carrying, suggesting that attempts to conceal deceit with non-verbal behaviour are largely unsuccessful. However, their study was very different contextually to that of interest in this thesis, so the same may not be true in an airport environment. At security checkpoints, with large, quick-moving crowds, time-consuming verbal measures which require some form of interaction may be impractical (Weinberger, 2010). Instances, when it is desirable to make an assessment discreetly, from a distance (Lansley et al., 2017), may also necessitate the use of non-verbal measures. Therefore, it is important to gain an understanding of how accurate people are when judging based on non-verbal behaviour alone and whether this form of assessment can be justified in this context.

This present study will assess accuracy using a virtual security checkpoint. Context can have an important effect on accuracy (Blair et al., 2010), and as this research is focused on deception in security contexts, virtual reality is used to depict this in a relatively easy, low expense manner (Pollard et al., 2020). The main questions that this study aims to investigate are whether in situations where some aspects of non-verbal behaviour are not significantly different between deceptive and honest people, how accurate are people at identifying deception from non-verbal behaviour and which factors do people rely on to detect deception? Ultimately, these results should contribute to determining whether assessments of non-verbal behaviour can be sufficiently accurate regardless of the idiosyncratic nature of non-verbal behaviour displays.

5.2 Methods

5.2.1 Participants

An a priori power analysis was conducted using G*Power software (Erdfelder et al., 2009) to determine the sample size that would allow the study to have sufficient power. The significance level was set at the standard $\alpha = .05$, the power level was set to $1\text{-}\beta = .80$, and a conservative effect size was set to .25. Accounting for potential data loss due to the use of virtual reality and participant withdrawal, a total sample size of 72 was determined to be sufficient. Seventy-two participants ($M_{age} = 20.38$ years, $SD_{age} = 2.23$ years) took part in this study. Within this sample, 66% were women, 32% were men and 2% identified as 'other'. Forty-six percent were British and the other 54% were non-British.

Participants were recruited by responding to posters placed around the Lancaster University campus, or via the university's online recruitment system for psychological studies. All participants were required to be at least eighteen years old, have normal or corrected to normal vision and have no previous adverse reactions to virtual reality. Depending on eligibility, they were either financially compensated or received course credit. Ethical approval was obtained from the Ministry of Defence Research and Ethics Committee and the Lancaster University Faculty of Science and Technology Ethics Committee.

5.2.2 Materials

Data collection took place in a standard laboratory, using an Oculus Rift CV1 HMD with a virtual reality compatible laptop and an Xbox controller.

5.2.2.1 *Virtual reality*

The virtual environment was similar to that used in the study in Chapter Four, except the participants' point of view was different (see Chapter Three, section 3.2). In this study, participants assumed the role of a security guard and were positioned so that they observed a range of avatars walking towards them.

5.2.2.2 *Avatars*

MakeHuman™ software was used to create the avatars whose non-verbal behaviour the participants were assessing. The avatars varied in age, race and gender. These superficial factors were identical for both the deceptive and honest avatars to try to mitigate the impact of these factors on decision-making (see Appendix A for an example of similar avatars).

5.2.2.3 *Avatar movement*

The non-verbal behaviour from participants in all the conditions in Chapter Four was used to animate the avatars in this study, using MVN Analyze™ and Unity™. The non-verbal behaviour from the participants in Chapter Four's package, weapon and future action conditions were used in the 'deceptive' condition in this chapter. This resulted in 12 avatars animated with honest non-verbal behaviour and 12 animated with deceptive non-verbal behaviour. This was to ensure that a wide range of non-verbal behaviour was used and should allow for a more comprehensive overview of people's ability to identify deception from different people's non-verbal behaviour. As mentioned in 3.2.1.2, there were four sets of the 24 avatars. Participants viewed only one of the four sets. Each set had physically identical avatars, which only differed in terms of the non-verbal behaviour that was used to animate the avatar. This was to ensure that accuracy was not impacted by any particularly unusual non-verbal behaviours.

5.2.3 Measures

5.2.3.1 *Accuracy*

Participants made assessed whether the avatars had deceptive intent via passive observation, reminiscent of actual security personnel in the SPOT program (Winter & Currier, 2015) using two buttons on the Xbox controller to communicate their decisions. Accuracy is defined as either the correct judgement of honesty or the correct judgement of deception.

5.2.3.2 *Post-experiment questionnaire*

A series of questionnaires were given, presented via Qualtrics software, which measured: virtual reality immersion (Appendix B), adapted from Usoh et al. (2000); participant demographics; which factors informed their assessment of the avatar's deceit, using two open-ended questions. All participants were required to provide comments about the factors that influenced their decision making, although the length of the comments varied between participants. The potential virtual reality immersion scores ranged from zero to 30.

5.2.4 Design

A repeated measures design was used, with the avatar's deception condition (honest or deceptive) as the independent variable. Accuracy was the dependent variable.

5.2.5 Procedure

Informed written consent was gained before the commencement of the study. Participants were given instructions detailing their role as a security guard which explained that they would view a series of avatars that would approach the security area and were required to decide whether they were honest or deceptive. They were then given a verbal explanation about the task and asked whether they had any further questions. The HMD was placed on their head, immersing them in the virtual environment. A series of 24 avatars approached the security area individually, allowing participants to observe their non-verbal behaviour. The order that the avatars were viewed was randomised for each participant. The participants were allowed as much time as they needed to view each avatar before deciding about whether they thought they were deceptive, using the controller. After they pressed a button on the controller, confirming their decision, a new avatar appeared. Once the study concluded, they completed the post-experiment questionnaires and were debriefed.

5.2.6 Analysis

5.2.6.1 *Accuracy*

Analysis was conducted using R studio. There were two questions that the analysis aimed to answer: 1) Does the overall decision accuracy differ from chance? 2) Does the avatar deceptive condition (honest or deceptive) and time spent viewing the avatars impact accuracy? A chi-squared test was used to answer the first question comparing the accuracy values with chance levels. A binomial logistic regression was conducted to answer the second question with the avatar deceptive condition as the independent (predictor) variable, time spent viewing the avatars as a covariate and decision accuracy as the dependent variable.

5.2.6.2 *Thematic*

An inductive thematic analysis was derived from the participant responses to the post-experiment questionnaire. The thematic analysis explores which factors informed participants' assessments of whether an avatar was honest or deceptive. NVivo software (QSR International, 1999) was used to code the participants' responses. These codes were reviewed to establish the common themes that existed amongst them, resulting in the six themes presented in section 5.3.2.

The thematic analysis discusses each theme and explains the relationship between the themes and how they were used to inform decision-making about whether an avatar was deceptive. Specifically, the thematic analysis responses refer to whether the avatar was *perceived* as deceptive, not whether the avatar was *actually* deceptive.

5.3 Results

5.3.1 Accuracy

Overall accuracy did not differ from chance $\chi^2(1, N = 72) = 1.22$, $p = .27$. Table 5.1 shows the results of the binomial logistic regression with the avatar deception condition as the independent (predictor) variable. Time spent viewing the avatars was included as a covariate.

Table 5.1 Binomial logistic regression results with avatar deceptive condition as the predictor variable (honest as the reference level), the time spent viewing the avatars as a covariate and accuracy as the dependent variable.

Predictor	Estimate	S.E	95% CIs		*z*	*p*
			Lower	Upper		
(Intercept)	0.38	0.15	0.10	0.68	2.63	<.01
Deception condition	-0.74	0.20	-1.15	-0.32	-3.55	<.001
Time	0.00	0.01	0.02	0.02	0.26	.79
Deceptive condition × Time	-0.01	0.01	-0.04	-0.02	-0.87	.39

Note. CIs = confidence intervals

Participants spent $M = 13.03$ seconds ($SD = 7.12$ seconds) viewing the avatars before making an assessment. Participants viewed the honest avatars for $M = 12.77$ seconds ($SD = 6.86$ seconds) and the deceptive avatars for $M = 13.28$ seconds ($SD = 7.37$ seconds). The overall accuracy rate was 49.31%, with the underlying distribution being 60.4 % correct classification of honesty and 38.2 % correct classification of deception. The virtual reality immersion scores in this study were quite poor ($M = 11.81$, $SD = 1.82$), a rate of 39%.

5.3.2 Thematic analysis

Six themes emerged from the data. The themes (physical appearance, disposition, walking behaviour, body positioning, looking behaviour and upper limb movement) capture different aspects of non-verbal behaviour that the participants used to inform their assessments of whether deception was present. The following thematic analysis provides an insight into how the themes were used.

5.3.2.1 *Physical appearance*

Some participants used the superficial characteristics of the avatars to inform their judgements of whether deceit was present. This use of superficial characteristics tracks with real-world occurrences of relying on such factors even though they are likely to result in biased outcomes (Homeland Security, 2002). Aspects of physical appearance, such as clothing were used to inform judgements of both those perceived as deceptive and honest. Whilst formal clothing was attributed with honesty: *"The ones in more official uniform such as suits looked more innocent",* casual clothing sometimes led to judgements of deceit: *"Those in tracksuit like clothes seemed to be more hostile"* (Participant 54).

As the ACLU reports, the TSA, often without justification, uses judgements of clothing such as "wearing improper attire" (ACLU, 2017, p.2) discriminately and inconsistently to help decide whether people are deceptive. Similar to the use of clothing, the extent to which an avatar was deemed as threatening was sometimes due to physical characteristics. Other physical factors such as gender and instances where non-verbal behaviour was perceived to imply a disability, were linked with judgements of honesty: *"Some form of disabilities led me to think they could be innocent"* (Participant 17). Given that the male avatars were no more likely to be deceptive than the female avatars, the supposition that: *"Men also seemed to be more hostile than the women"*(Participant 54) suggests that some salient physical characteristics may detrimentally influence perceptions of whether deceit is present.

5.3.2.2 *Disposition*

Disposition was often cited as influencing decision-making, with traits such as nervousness ascribed to both avatars perceived as honest and deceptive. Participant 33 commented that: *"A lot of movements which I would call nervous led me to believe people had hostile intent. I think the hardest decision was to decide if the NPC [non-player character] was nervous or had hostile intent"* which reflects a general issue of distinguishing between appropriate situational nervousness induced by the security environment and nervousness influenced by deception. Although, some people claim that working in security environments allows for the development of a "sixth sense" for detecting deceit related nervousness (Pinizzotto et al., 2004, p. 6), realistically this is a subjective ability, which is questionable in its reliability (Wigginton et al., 2014). The nervousness aspect of disposition was also cited as an influencing factor in determining honesty: "Nervous innocent scratching of the face or arms when talking at the check-in desk" (Participant 29). This interpretation of nervousness, along with the previous quote, exemplifies the subjective nature of aspects of disposition and how they are perceived by others. It also raises the question of whether some people are better able to distinguish between regular nervousness and deceptive nervousness, as the above quote implies that the nervousness was not reflective of deception.

Perceived confidence, another aspect of disposition, did not appear to have such conflicting perception compared to nervousness, as it was cited as influencing decisions of honesty such as: "When they walked confidently up to the desk without much hesitation and when they didn't fidget" (Participant 77).

5.3.2.3 *Walking behaviour*

The way that the avatars walked was an attribute that participants frequently mentioned as having informed decision-making. Overlap between themes was notable with walking behaviour. Specifically, aspects of disposition were often reported to manifest through the style of walking: “Walked up to the desk with confidence, good posture, no hesitation” (Participant 45). The following comment highlights the observation of a potential link between disposition and walking behaviour, suggesting that nervousness may have resulted in slower walking behaviour, which was then perceived as reflecting honesty: “There were times when they did walk slowly or held [their] arms in an odd way, but sometimes this seemed to present itself as nerves instead of hostile intent, so I believed that they were innocent” (Participant 7).

The same behaviour, walking slowly, was at times interpreted in a contrasting way to the above quote. Walking paces that appeared to deviate from the norm were perceived as being reflective of deception, such as when a mediating factor such as presumed innocent nerves, or other, was not detected to explain an unusual walking pattern: “If the person moved slower or faster than normal it made me believe that they were guilty of something” (Participant 36). This contrasting interpretation of a slow walk in the previous quote again conveys the potential subjectivity of assessments of non-verbal behaviour, a phenomenon which is also evident in the real world according to reports of subjective profiling in airport security (ACLU, 2017; Baker, 2002; Hashad, 2004; Meyer, 2010;).

5.3.2.4 *Body positioning*

Body positioning revolves around how the avatars were positioned, both in terms of posture and relative to their proximity to the security checkpoint and staff. Avatar proximity varied because participants in Chapter Four varied to a degree in how close they chose to stand next to the check-in desk whilst being interviewed. An absence of movements outside of the norm was noted as influencing decisions of honesty: “Nothing [seemed] suspicious, they just walked towards to security checkpoint without too much movement or too little movement” (Participant 76). In keeping with behaving within norms, maintaining “ a medium distance” (Participant 51) from the staff was cited as influencing judgments that they were honest.

5.3.2.5 *Looking behaviour*

Looking behaviour refers to the eye and head movements of the avatars. Similar to the body positioning theme, an excess or lack of eye contact and looking around was reported as influencing perceptions of deceit. However, the subjectiveness of how looking behaviour was interpreted, and the comfortability of sustained eye contact was quite evident: “Any non-eye contact in my eyes is seen as more suspicious” (Participant 41). Conflicting with the previous quote where eye contact appears to be encouraged to convey honesty, the following shows the opposite effect: “A few of them stared at me, that seemed quite hostile” (Participant 77).

5.3.2.6 *Upper limb movement*

Upper limb movement refers to the arms and hands, which were mentioned as being influential to decision-making This theme potentially has some overlap with disposition, for example, avatars with their “hands clearly visible” (Participant 60), were interpreted by some as literally suggesting that they had nothing to hide, thus they must be honest. In contrast, hands that were not visible were reported as inducing decisions of deception. This could be interpreted to suggest that there is an underlying assumption that the hands are a source of ‘leakage’ of deceptive behaviour (Sartori et al., 2016) and resultingly, trying to mask deception potentially manifests physically in concealed hand movements, for example: “Some of them were hiding their hands in their pockets, which could mean they were masking their feelings or were afraid their hands would start shaking" (Participant 35).

5.4 Discussion

5.4.1 Accuracy

The results show that overall accuracy rates were low (49.31%) and no different from chance. The underlying distribution of this accuracy was 60.4% correct identification of honesty and poor 38.2% identification of deception. Six themes emerged from the thematic analysis: physical appearance, disposition, walking behaviour, body positioning, looking behaviour and upper limb movement, providing an insight into which factors may influence decision-making in this context.

The poor decision accuracy in this chapter suggests that assessments of non-verbal behaviour in instances where non-verbal behaviour may not be significantly different between honest and deceptive people (as could be the case for some sections of the population (Blackwood et al., 2015) in the real world) may not be sufficiently accurate. The accuracy rates found in this study were poor, no better than chance and in line with those reported in the literature, (Bond & DePaulo, 2006). The poor accuracy rates within this controlled laboratory environment may suggest that in a real airport where there are other distracting factors, assessments based on non-verbal behaviours may be hindered even more.

However, the study's poor overall accuracy rates (Bond & DePaulo, 2006) may be due to the relatively poor virtual reality immersion rates. Blair et al., (2010) suggest that heightening contextual relevancy of experiments should lead to improved accuracy, thus the poor virtual reality immersion may have impacted the decision accuracy results. There are a few reasons why the immersion was poor in this study, especially in comparison to Chapter Four where the immersion was 33% higher. The virtual experience was not as interactive compared to Chapter Four, because the participants did not speak with the avatars, they merely observed their behaviour. Also, the participants in this study did not walk around in the virtual environment as they did in Chapter Four, meaning that fewer senses were stimulated, which can lead to poorer immersion (Limbasiya, 2018), though could hear airport audio stimuli. This spurs the question for future research of whether better accuracy rates can be obtained in this context with a similar paradigm that incorporates more stimulation of the senses to gain better immersion in virtual reality whilst participants are assessing others. Incorporating a virtual environment more like mixed reality, as was the case in Chapter Four, may be beneficial for future research. A more immersive virtual environment could include walking around whilst in virtual reality and correspondingly in the real world within the laboratory space. Interacting with physical objects, such as a security desk, in the virtual and real-world could also help.

5.4.2 Thematic analysis

The way that some aspects of the themes were interpreted and used to assess deception mirrored the critiques of some airport security protocols (Homeland Security, 2002; ACLU, 2017), namely regarding their subjectivity. The perception of aspects of disposition, such as nervousness, resulted in varying and often contradictory rationales for presumptions of deceit or honesty. Concerning the perceptual experience hypothesis (Cañal-Bruland et al., 2010), personal experience with nervousness may influence how it is perceived in others. For example, someone who experiences nervousness in certain environments may recognise similar nervousness in others as benign, whereas a hyper-confident person may interpret it as more sinister due to the lack of personal experience with nervousness in that context. Another aspect of the interpretation of disposition that is reminiscent of EVT was that some participants used an absence of a trait, for example, hesitant and fidgety behaviour, to rule out deception. Using disposition to influence decisions could be problematic given the unfounded links with actual deception and the subjectivity of perception.

Likewise, the interpretations of the looking behaviour theme draw parallels to the Expectancy Violations Theory (Burgoon et al., 2005) since both too much and too little eye contact appeared to violate behavioural norms, whilst also emphasizing the subjectivity of how people interpret eye contact. Similarly, normal body positioning did not violate expectations and was perceived by some participants as reflecting honesty. However, this theme could be inferred to display the potential for innocuous abnormal positioning to be falsely identified as perceiving deception. These decision-making rationales are evidence of how in a more general sense, the inclusion of broad subjective traits as markers of deception in security protocols may potentially lead to problematic usage in practice (Meyer, 2010) since objectively defining an acceptable occurrence of these aspects of behaviour is difficult and perceptions of what violates the norm may vary between individuals.

The thematic findings provide a different insight into how people form their judgements, which is important if we are to try to reduce the reliance on invalid factors to try to gain an understanding of this type of decision-making. One question that emerged from the thematic analysis is whether the reliance on subjective factors is the main factor contributing to poor accuracy or whether non-verbal behaviour is unsuitable for deception detection regardless? The following chapter addresses this query by incorporating two techniques of information reduction to try to limit the influence of subjective, biasing factors.

5.4.3 Limitations

It is important to question the extent to which the decision task in this chapter is reflective of the role that security agents are tasked with in the real world. Visitors to high-security environments such as airports are constantly assessed about whether they pose a security risk. For example, Israel’s Ben Gurion airport has at least six layers of security, which begins with a roadside check, before passengers even step foot in the airport itself (Kelly, 2009). The security personnel at Ben Gurion are trained in behaviour profiling and ask simple questions to determine if people pose a threat whilst looking for behavioural indicators that suggest nervousness or distress (Kelly, 2009). Likewise, the SPOT program involves observing a range of non-verbal behaviours and determining whether the amalgamation of these behaviours exceeds a threshold of honest behaviour (Winter & Currier, 2015). As such the decision task in this chapter was devised to reflect the overarching concept of observational behavioural profiling, albeit simplistically, without prior instruction on which cues to look for and within the confines of a laboratory. Consequently, although these results provide an insight into decision-making, caution should be applied in the attribution of these findings to real security checkpoints, where behavioural profiling use varies between countries and/or cities (Wigginton et al., 2014) and may be more complex.

The sample population of this chapter, whilst diverse in nationality and gender, did not reflect a wide age range, nor occupational diversity. The bulk of participants were young university students, as is common in research, and consequently, their own travel experience (and exposure to security checkpoints) may have been limited and their experiences were presumably solely from the point of view of a passenger. It would be a good idea for future research to consider a similar procedure with a mix of laypeople and highly trained security personnel to compare the differences in the factors used to aid decision-making.

5.4.4 Summary

In conclusion, this study illustrates that people are poor detection detectors overall and that assessments based on non-verbal behaviour, particularly in instances where non-verbal behaviour is not distinct regardless of deception, may not be a suitable practice. The thematic analysis sheds some light on the factors that are used to influence decision-making and shows that there are differences in whether people perceive deception, depending on how they interpret different aspects of non-verbal behaviour. The following chapter will investigate whether the accuracy rates of detecting deception found in this chapter can be improved by incorporating two strategies to reduce the salience of irrelevant information.

Chapter Six - Information Reduction Techniques and Decision Accuracy

6.1 Introduction

The absence of non-verbal cues that distinguished honesty from deception, as in Chapter Four, coupled with the unreliable nature of non-verbal cues in the literature, the subjective interpretation of non-verbal behaviour (Chapter Five) and the critiques about the use of ethnicity, race, gender and religion (ACLU, 2017) by security personnel, lean towards assessments of deception based on non-verbal measures being unsuitable.

To further contribute towards providing an overall assessment of the appropriateness of using non-verbal behaviour measures, this chapter will determine whether lessening the salience of potentially biasing factors can result in the identification of deception above chance levels, even in cases where non-verbal behaviour does not differ between honest and deceptive people. The rationale for this chapter has emerged from the poor accuracy found in Chapter Five and the literature and the reported use of superficial factors to inform decision-making in Chapter Five. This chapter aims to investigate whether the poor accuracy rates can be improved to a rate that warrants the use of non-verbal measures, or whether it is merely a subjective guessing game. This chapter will look at two techniques of information reduction (and accordingly, the two methods of decision-making, non-conscious and conscious, that underpin them) to reduce the ease of relying on superficial factors that may prejudice judgements. Other latent factors may come to the fore and aid decision-making using these techniques. Whether either method of information reduction can improve the accuracy of detecting deception beyond chance will be determined. First, the *amount* of information present will be reduced by imposing constraints on the amount of time that the visual information will be available, potentially leading to judgements that are based on intuition. Second, the *type* of information present will be manipulated with the use of point-light displays (PLDs), which reduce the presence of superficial information, potentially ensuring that judgements are predominantly uninfluenced by superficial factors.

Gestalt theory, when applied to this research, would suggest that looking at people as a whole is better than decision-making based on smaller (potential biasing and difficult to interpret) components. Within decision-making tasks, more salient factors, such as ethnicity, clothing and gender may detrimentally bias people's assessments as suggested in the thematic analysis in Chapter Five and the reports of systematic bias within security personnel (ACLU, 2017; Committee on Science, Space and Technology, 2011; Meyer, 2010; Winter & Currier, 2015). A "loss of individualization" of people that are perceived to have an increased likelihood of being deceptive (Harcourt, 2006, p. 109), reflected by applying intense scrutiny at airports to specific populations, can hinder deception detection rates by producing an inflated belief that the remaining population is unlikely to commit criminal acts (Press, 2010). Therefore, stripping back the prominence of some superficial factors may be of interest due to reducing the propensity to focus on small factors and provide a more gestalt assessment.

Limiting the amount of time that information is visible could be a promising method of improving accuracy in deception detection research. The unconscious thought principle posits that there are two aspects of processing, conscious and unconscious (Dijksterhuis & Nordgren, 2006). Conscious processing has a limited capacity, constrained by memory and executive functioning processes (Baddeley & Hitch, 1974; Kahneman, 1991). In contrast, unconscious processing does not have these limitations (Dijksterhuis & Nordgren, 2006) and could be a promising strategy for increasing decision accuracy above chance if induced in decision-making procedures. Thin slicing is a method of reducing the amount of time that information (verbal, auditory, visual) is available, which can allow for the recognition of behavioural patterns via unconscious processing that are otherwise not consciously perceptible. Reducing snippets of behaviour to as little as 5 seconds is thought to allow for the accurate perception of information such as “personality, affect and interpersonal relations” (Ambady, 2010, p.271) compared to observing longer instances of behaviour. Reducing stimuli to thin slices of behaviour may coerce unconscious (automatic) processing because it encourages snap judgements which may be informed by factors that cannot be verbalised/tangibly perceived. In contrast, exposure to longer instances of behaviour may encourage more conscious, deliberate processing and allow irrelevant but salient factors to influence decisions.

It is reasoned that people are unconsciously able to identify “macro traits”, which could be likened to gestalt-like assessments (Thompson, 2012), through thin-slicing, such as trustworthiness and nervousness. In contrast, the “micro traits”, which reflect more individual bodily cues (Thompson, 2012) such as eye contact and hand gestures require conscious processing. Although the micro traits may map onto the macro traits, decision-making based on the micro traits may be more difficult because it relies on conscious processing of how and why reduced hand gestures, for example, could indicate deception. Consequently, unconsciously processing macro traits may benefit deception detection, in line with gestalt theory, by heightening people’s propensity for their judgements to be influenced by the person in question as a whole, rather than homing in on individual and potentially misleading small factors. Hypothetically, methods that encourage judgements based on unconscious processing, such as thin-slicing (Ambady, 2010), may lead to better accuracy. It is possible that limiting the amount of time that people have to view the visual stimuli may result in unconscious processing coming to the forefront in decision-making and better accuracy at detecting deception (Albrechtsen et al., 2009). Ambady (2010) found that imposing time constraints was an effective method of inducing unconscious processing which can lead to more accurate judgements.

Eyewitness testimony is an example within forensic psychology of how time can be linked with better accuracy. The identification of suspects was more accurate when less time was spent deliberating on suspects in a line-up (Gustafsson et al., 2019; Lindholm et al., 2018), suggesting a relationship between accuracy and time provided to process information. In the aforementioned scenario, it could be reasoned that longer deliberation and visibility of the information (suspects in this instance), is reflective of doubt. The same effect may be true, relative to deception detection. Could accuracy of detecting deceit in the literature (Bond & DePaulo, 2006) be hindered by unconstrained visual availability of the stimuli? There have been conflicting results surrounding this notion in deception detection. Research has sought to induce unconscious processing in two ways. Reinhard et al. (2013) first showed participants a video for which they had judge for deceptive traits and then induced unconscious processing by engaging participants with a crossword for three minutes before they had to make their judgements. Albrechtsen et al. (2009) induced unconscious processing by limiting the length of the videos that participants had to judge. It is important to note that both Reinhard and Albrechtsen's methods aim to exploit the same construct, unconscious processing, in different ways.

The study in this chapter builds on Albrechtsen's time-constrained method of inducing unconscious processing to see whether coercing more gestalt-like processing by limiting the duration of the visibility of the stimuli (and reducing the likelihood that people focus on perceived individual cues of deception) positively impacts decision accuracy compared to chance. Albrechtsen et al. (2009) found that limiting the visibility of stimuli to 15 seconds (comprised of three 5 second clips which were joined together) was effective at improving accuracy. However, 15 seconds is a long time, in which superficial factors can be observed and potentially impact decision-making. This study aimed to see whether reducing this time further by two-thirds could be effective in reducing the salience of superficial characteristics, with the need to minimise the duration of security measures in mind (Lee & Jacobson, 2011).

Albrechtsen et al. (2009) and Reinhard et al. (2013) both found that inducing unconscious processing leads to improved accuracy compared to more deliberative processes. However, others have not found any improvement in accuracy when inducing unconscious processing (Moi & Shanks, 2015; Wu et al., 2019). This chapter aims to address the conflicting literature on unconscious processing by applying this concept to the methodology used in Chapter Five, to see whether improved accuracy of detecting deception can be gained by limiting the amount of time that participants can view the avatars in virtual reality.

Bias, both unconscious and conscious, is a part of human nature (Cueller, 2017; Fuentes, 2016). When judging the presence of deceit, our decisions may be influenced by factors other than those which are relevant to deceit itself, such as gender, ethnicity and religion (Baker, 2002). For example, the TSA has been condemned for disproportionately targeting men who are perceived as Arab or Muslim (ACLU, 2017) and within interrogation contexts, suspects are less likely to be perceived as deceptive if they are deemed to be attractive (Murai et al., 2018). Superficial factors such as perceived ethnicity, religion and attractiveness can influence perceptions of deception and have a potentially detrimental effect on decision-making. The presence of these irrelevant characteristics may increase the likelihood of making a decision based on a biasing characteristic (Murai et al., 2018). Within security contexts, the salience of some characteristics may be problematic due to the potential for discriminatory practices since profiling in this manner at security checkpoints is both legally and ethically questionable (Hurrell, 2013; Zedner, 2013). As a result of the influence of bias, deceit may be wrongly overlooked in those who do not conform to a particular stereotype (Meyer, 2010), and undeserved scrutiny may be given to those who do.

Point-light displays are one method of reducing superficial information in non-verbal behaviour because attributes such as gender, religion and ethnicity are absent or less salient. Research has used point-light displays to induce good accuracy rates of detecting deception (Runeson & Frykholm, 1983). A series of tasks where participants judged point-light displays involved people who attempted to falsely convey the weight of a box they were carrying and another task involved people who attempted to conceal their identified gender (Runeson & Frykholm, 1983). Participants were able to identify deceit in up to 83% and 75% of cases respectively. Despite attempts at behavioural control to conceal their deceptive intentions, the honest action (box weight) and state of being (gender) were still discernible from point-light displays.

Similarly, in a sports psychology study, the presence of superficial information was reduced by manipulating the spatial frequency of video stimuli. Park et al. (2019) showed participants a series of videos of badminton shots. The deceptive badminton shots involved the players using misleading eye gaze or head movements so that the shuttlecock landed in an unpredictable part of the court. The video did not show the frame in which the racket hit the shuttlecock to make participants' judgements of whether the badminton player was being deceptive more difficult. Lower spatial frequency caused the videos to be blurrier, reducing the visibility of the superficial features, compared to the high spatial frequency videos in which superficial features were salient. People were better at identifying deceptive shots when the spatial frequency was low (reduced superficial information), compared to normal and high spatial frequency videos, where the superficial information was more visible (Park et al., 2019).

Both point-light displays and spatial frequency manipulation as a method of reducing superficial information seem promising in their effect on accuracy. However, both Park et al. (2019) and Runeson and Frykholm (1983) focused on attempts to deceive by manipulating physical movements e.g. walking differently to appear as the opposite gender or misleading head movements before taking a badminton shot. In contrast, to contribute to security checkpoint research it is pertinent to investigate whether reducing superficial information is beneficial to detecting attempts at cognitive deception, such as concealed intent more so than deception due to deliberate manipulation of bodily movements. Specifically, if people are deceptive about being in possession of a concealed weapon or intending to commit an illegal act in the near future, can deception be identified, above chance rates even in the absence of reliable/valid cues, by making superficial variables less salient? Furthermore, the context of the aforementioned studies, badminton and identifying the true weight of a box or person's gender, makes it difficult to deduce whether similar effects of reducing superficial information may be gained within a virtual security context.

In recent years there has been some concern, from civil liberties groups and individuals, that there may be an acceptance within the industry, post 9/11 (Baker, 2002), to profile people based on characteristics such as ethnicity, race and religion to determine whether they deserve extra scrutiny (Meyer, 2010; ACLU, 2017). The use of profiling in this way has been challenged as being both inherently flawed and potentially discriminatory (Aviation Security, 2002).

Ultimately, this chapter aims to answer one main question. Can techniques to de-bias decision-making bolster accuracy to a level that warrants its use in high-security environments, even considering the reliability and validity issues of non-verbal cues? Point-light displays reduce the salience of superficial factors and may be easily introduced when viewing live surveillance videos, resulting in scenarios where the decision-maker views the point-light displays of passengers on a computer screen, without seeing the actual person, thus reducing the likelihood of any biases influencing their judgements. Likewise, time constraints could be introduced if judgements are made by watching people through a computer with time limits. As well as comparing the two information reduction conditions to each other, the accuracy in the time-constrained condition and reduced superficial information condition will be compared to chance levels to see if either information reduction condition is more beneficial in aiding accuracy.

6.2 Methods

6.2.1 Participants

An a priori power analysis was conducted using G*Power software (Erdfelder et al., 2009) to determine the sample size that would allow the study to have sufficient power. The significance level was set at the standard $\alpha = .05$, the power level was set to $1-\beta = .80$, and a conservative effect size was set to .25. Accounting for potential data loss due to the use of both virtual reality and point-light display technology, as well as participant withdrawal, a total sample size of 54 was determined to be sufficient. Fifty-four participants (78% women, 22% men, $M_{age} = 20.24$ years, $SD_{age} = 2.82$ years) were included in this analysis. Within this sample, 54% were British and 46% were non-British. Three participants were excluded from the original sample.

Participants were recruited by responding to posters placed around the Lancaster University campus, or via the university's online recruitment system for psychological studies.

All participants were required to be at least eighteen years old, have normal or corrected to normal vision and have no previous adverse reactions to virtual reality. Depending on eligibility, they were either financially compensated or received course credit. Ethical approval was obtained from the Ministry of Defence Research and Ethics Committee and the Lancaster University Faculty of Science and Technology Ethics Committee.

6.2.2 Materials

Data collection took place in a standard laboratory, using a virtual reality compatible laptop, Oculus Rift CV1 HMD (time-constrained condition only) and an Xbox controller.

6.2.2.1 *Virtual reality*

The virtual environment, created using Unity™, was similar to that used in Chapter Five (see Chapter Three, section 3.2 for details). This consisted of a checkpoint environment, with a range of approaching a security area. The non-verbal movement gained from participants in Chapter Four was used to animate the avatars in this study. The avatars varied in age, race and gender avatars (see Appendix A). These factors were similar for both honest and deceptive avatars.

6.2.2.2 *Point-light displays*

The non-verbal data gained from Chapter Four was used for the point-light displays (see Chapter Three, section 3.3) in this chapter. The point-light displays were created using PLAViMoP software (Decatoire et al., 2018) and were presented to participants on a laptop, using a 2D version of Unity, which does not require a head-mounted display.

6.2.3 Measures

6.2.3.1 *Accuracy*

Accuracy is defined as either the correct judgement of honesty or the correct judgement of deception. Participants used two buttons on the Xbox controller to communicate their assessment of the avatars as either honest or deceptive.

6.2.3.2 *Post-experiment questionnaire*

Questionnaires were given, which measured virtual reality immersion, adapted from Usoh et al. (2000) (Appendix B) and general participant demographics. The potential immersion scores ranged from zero to 30.

6.2.4 Design

This study used a repeated measures design, with the information reduction condition: time constraint (virtual reality) and reduced superficial information (point-light display) as the first independent variable and the deception condition (honest or deceptive non-verbal behaviour) as the second independent variable. The condition order was counterbalanced. The deception condition refers to the movement animating the avatars or point-light displays, gained from the non-verbal data of participants in the study in Chapter Four. This design was chosen to reduce the potential effect of participant variance on accuracy performance and to gain a better insight into the difference in an individuals' accuracy between the two information reduction conditions.

6.2.4.1 *Experimental manipulation*

Virtual reality was used for the time-constrained condition. The participant role and task were the same as in Chapter Five. Participants were positioned in the virtual environment so that the avatars were walking towards them. Participants in Chapter Five were not restricted in the amount of time that they could view the avatars before decision-making. They viewed the avatars for $M = 13.03$ seconds, $SD = 7.12$ seconds. To see whether deliberately imposing restrictions (unlike Chapter Five) would benefit accuracy relative to chance, this study imposed a limit on the presentation of each avatar. Each avatar was shown for a maximum of 5 seconds before a screen appeared, blocking the participants' view of the security checkpoint environment. The imposed time restriction forced participants to communicate their decision of whether the avatar was deceptive as quickly as possible, using the Xbox controller.

The reduced superficial information condition was conducted in 2D because previous studies have shown that people are good at extracting information from point-light displays in the 2D format. For example emotions (Dittrich et al., 1996), the direction of movement (Kuhlmeier et al., 2010) and deceit (Runeson & Frykholm, 1983) were all identifiable using 2D point-light displays. Also, the point-light display condition is comprised of a completely black background/scene, with just the white dots visible, which was found to be unsettling during pilot testing in virtual reality. Due to this reason and the previous successful use of point-light displays in 2D, conducting this study in 3D was not deemed to be more beneficial in potentially aiding deception detection in comparison to potential effects on participants' performance.

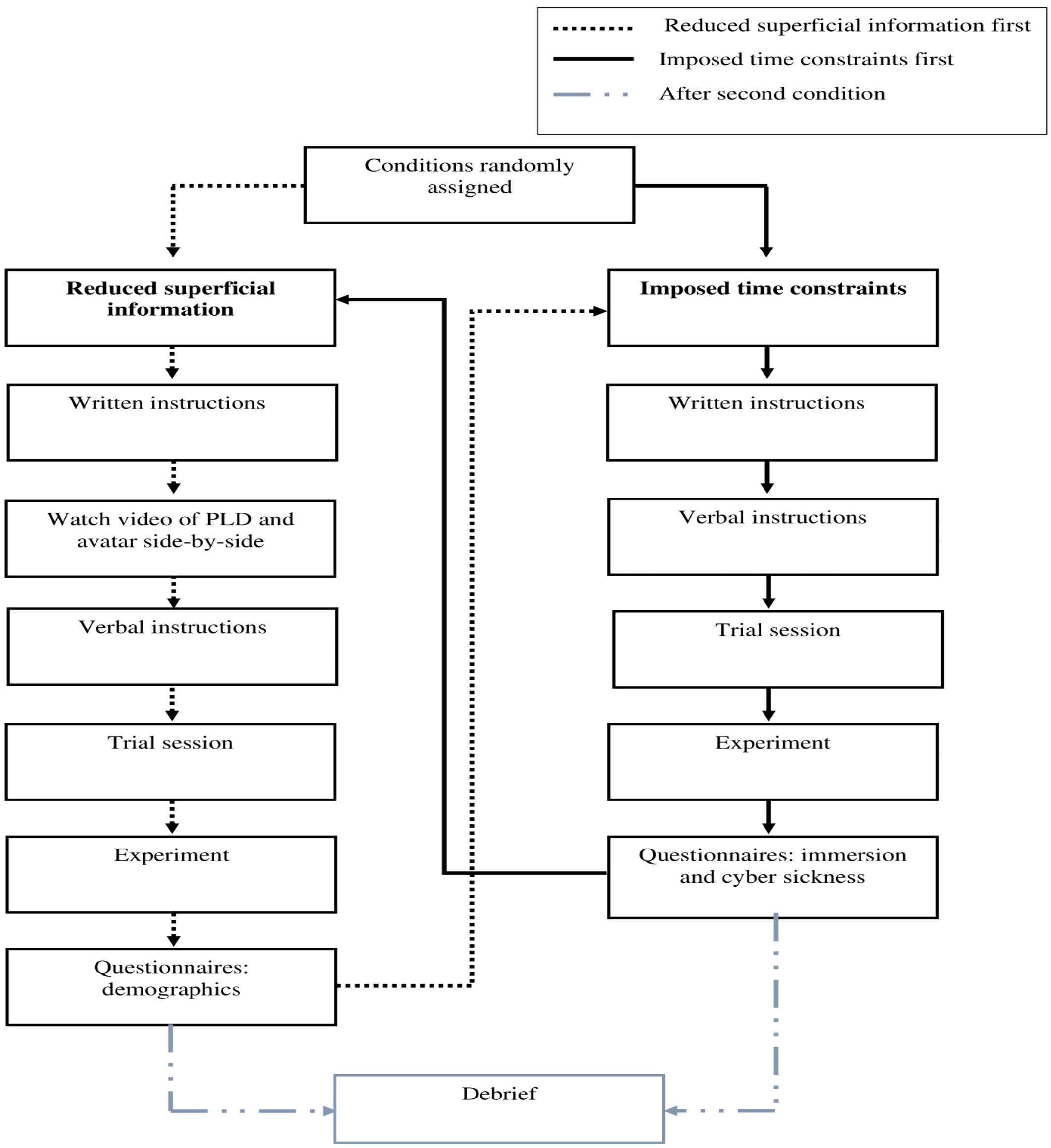
6.2.5 Procedure

Informed written consent was gained before the commencement of the study. Participants were seated in front of a laptop for both conditions. The information reduction condition order was randomly assigned so that half of the participants completed the time-constrained condition first, followed by the reduced superficial information condition, and vice versa. As such, the following description is from the perspective of completing the time-constrained condition first followed by the reduced superficial information (see Figure 6.1 for an illustration of the study flow).

Instructions were given to participants, detailing the task. Participants were instructed that they were assuming the role of a security guard and would be judging a series of avatars that would the security area. They were informed that they would have only a limited amount of time to view each avatar before their vision of the virtual scene would be occluded, requiring them to decide whether the avatar was honest or deceptive, as quickly as possible, using the Xbox controller. Following this, they were given the opportunity to ask any questions. Then they had a trial session to allow them to become accustomed to the time constraints. The time-constrained task consisted of 30 avatars in random order. Each avatar was visible for 5 seconds, after which a screen appeared prompting a quick judgement of whether the avatar was honest or deceptive. Once the time-constrained condition concluded, the immersion questionnaire (Appendix B) was provided. Following the questionnaire, participants were offered a break before starting the reduced superficial information condition.

Instructions of the reduced superficial information task were given. Like the time-constrained instructions, participants were informed that they were to assume the role of a security guard and assess whether the point-light displays that walked towards the screen were either deceptive or honest. This was followed by the opportunity to ask questions. They were shown a short video of a point-light display in motion next to an avatar, before starting a short trial session. Point-light displays were novel for most participants, so this allowed them to understand how the point-light displays' movement mapped onto regular/avatar movement. This task required the participant to assess 30 randomised point-light displays. Participants communicated their decision of whether the point-light displays were deceptive using the Xbox controller. Once this part of the study concluded, the demographics questionnaire was given and then the participants were debriefed.

Figure 6.1 Study flow with both the reduced superficial information and imposed time constraints condition



6.2.6 Analysis

R Studio was used for the analysis to answer three questions. 1) Does the imposition of time constraints improve accuracy compared to chance? 2) Does reducing superficial information improve accuracy compared to chance? 3) Is one technique (imposed time constraints or reduced superficial information) superior to the other? To answer the first and second research question a chi-square test was used to compare the observed accuracies with chance. A logistic regression was used to answer the third research question, with the information reduction condition (reduced superficial information and time constraints) and deception condition (honest and deceptive avatars/point-light displays) as the independent (predictor) variables. Accuracy was the dependent variable. The participants and the individual avatars varying among participants were specified as random effects.

6.3 Results

There was no significant difference in accuracy between the time-restricted condition and chance, $\chi^2(1, N = 54) = 3.21$, $p = .07$. There was a significant difference in accuracy between the reduced superficial information condition and chance, $\chi^2(1, N = 54) = 8.91$, $p <.01$, with less accuracy compared to chance.

Table 6.1 Results of the logistic regression, with accuracy as the dependent variable

Predictor	Estimate	S.E	95% CI		*z*	*p*
			Lower	Upper		
(Intercept)	0.22	0.06	0.10	0.35	3.46	<.001
Information reduction condition (time- constraints)	0.04	0.09	-0.14	0.22	0.41	0.68
Deception condition (deceptive)	-0.79	0.10	-0.99	-0.58	-7.56	<.001
Information reduction condition × deception condition	-0.24	0.15	-0.53	0.05	-159	.11

Note. Information reduction condition (point light as reference level), deception condition (honest as reference level), and the information reduction condition × deception condition interaction are listed as predictors.

Overall accuracy was not significantly different between the information reduction conditions (see Table 6.1): reduced superficial information (47.8%), imposed time constraints (46.6%). There was no significant interaction between the information reduction conditions and deception condition on accuracy (see Table 6.1); reduced superficial information: honest (55.6%), deceptive (36.3%); time-constrained: honest (56.5%), deceptive (31.8%).

6.4 Discussion

There was no benefit to overall accuracy compared to chance levels of either the reduced superficial information condition (47.8%) or the time-constrained condition (46.6%). The underlying distributions of accuracy within both information reduction conditions support the truth default theory (Levine, 2014), whereby people assumed honesty more than deception. In the reduced superficial information condition, correct identification of honesty was 55.6% compared to identification of deceit, 36.3%. Likewise, in the time-constrained condition, correct identification of honesty was 56.5%, compared to deceit 31.8%.

The lack of a benefit to overall accuracy with either information reduction conditions contrasts with the literature which would suggest that the reduced information should have elicited better accuracy rates, thereby also conflicting with previous suggestions that findings of abilities to detect deceit apply to different contexts (Frank & Ekman, 1997). One reason for the time-constrained conditions' contradictory findings with studies that have found that unconscious processing improves accuracy (Albrechtsen et al., 2009; Reinhard et al., 2013), may be that the incorporation of unconscious processing with virtual reality was too overwhelming. Virtual reality was novel to a lot of the participants. Coupled with the time constraints, the lack of a difference compared to chance may be due to a combination of the limited time and the novelty and potential sensory overload of the task. Given the contrasting results within the literature on unconscious processing, relative to deception detection, future studies may wish to introduce time constraints when judging the presence of deception in video clips of deception at a security checkpoint (Abernethy et al., 2010), as opposed to virtual reality, to reduce the number of novel stimuli/methodologies.

However, other researchers have also struggled to find positive effects when eliciting unconscious thought (Moi & Shanks, 2015; Wu et al., 2019). Albrechtsen et al.'s, (2009) study improved accuracy using 15-second clips to elicit intuitive (unconscious) processing, whilst this study reduced that time by two-thirds. Given the rationale behind thin slice judgements and accuracy (Ambady, 2010) it was presumed that reducing the viewing time further to 5 seconds, as this study did, would have been beneficial to accuracy compared to chance. Therefore, the findings from this chapter spur the questions of whether thin-slicing generally is an appropriate tool for inducing unconscious processing in this experimental context and whether thin-slicing to the extremes used in this study (five seconds) can be beneficial? The wider difficulty to replicate Reinhard's unconscious decision-making findings (Street & Vadillo, 2016) and the effect of time-constraints on accuracy (Albrechtsen et al., 2009) may indicate that a Type I error resulting from random assignment/sampling or environmental differences contributed to their results, rather than unconscious processing alone. The other null findings in this area may suggest that perhaps inducing unconscious processing is simply not beneficial to deception detection.

The demographic of the sample may have influenced the results in this study, explaining some of the conflicts with the literature. Park et al. (2019) found that badminton experts had better accuracy rates when superficial information was reduced, compared to novices. It may be inferred that the benefit of reducing superficial behaviour is contingent on contextual familiarity.

Consequently, the lack of difference between information reduction conditions and chance may have been due to the sample population, as the participants were predominantly students. A potential future modification of this study may include a sample consisting of experts who work on the frontlines in the security sector, to see if improvements in accuracy can be gained.

Moreover, it may be worthwhile to assess in the future whether incorporating a prisoner or experienced criminal population would lead to a benefit of reducing superficial information. Though this would not directly benefit security, it would contribute to developing our understanding of what leads to good deception detection and potential training methods for improving that skill. Some have attributed the ability to detect deception in point-light studies to activation of the mirror neuron system (Pinar Saygin et al., 2004), which suggests that the same areas of the brain are activated when both performing an action and observing an action. This could explain the finding that prisoners interpretation of non-verbal cues to deception were more accurate than both other experts (officers, detectives and prison guards) and laypeople (Vrij & Semin, 1996). Their increased experience of performing deceptive actions themselves (Bond, 2008) may enhance their knowledge of how behaviour changes when people are deceptive. Reduced superficial information may be worth further exploration as other studies have found positive effects on accuracy. At the very least, assessing if other populations perform better than this predominantly student sample could be beneficial for learning about what contributes to better than chance levels of accuracy at detecting deception.

6.4.1 Summary

In conclusion, the findings of this study do not offer evidence that either form of information reduction is a useful method of improving the accuracy of non-verbal deception detection above chance levels.

Part IV: Chapter Seven – Gaze Behaviour Before Deception Identification

7.1 Introduction

This final empirical chapter builds on Chapters Four and Five. Chapter Four focused on which cues people exhibit when deceptive, whilst Chapter Five focused on people's ability to judge the presence of deceit in others and provided a self-reported insight into what informed their judgements. These two chapters have revealed that the non-verbal behaviours were assessed did not differ between those who were honest and deceptive (Chapter Four), and that decision-making was inaccurate and informed by subjective perceptions of some factors (Chapter Five). This present chapter is exploratory and aims to build on Chapter Five by providing a more objective insight into people's decision-making. This chapter uses eye-tracking technology to assess which areas of the body people *actually* focus on prior to making judgements and the relationship between focusing on particular parts of the body and accuracy.

7.2 Eye-tracking and deception research

Brunswik's lens model (Brunswik, 1952), when applied to deception research, emphasises the need to explore deception detection from both the point of view of the person committing the act of deception (Chapter Four) and the person tasked with detecting deception (Chapters Five and Six). Although the three previous empirical chapters have approached deception detection from both sides of the lens, it is important to take into consideration that there may be discrepancies between what people think has informed their decision-making and what they actually looked at (Hartwig & Bond, 2011). This difference may be due to the limitations of people's meta-cognition (Lawson et al., 2013), which may impact what is reported as having influenced decision making, as in Chapter Five. Although we are constantly looking at the world around us, it is too cognitively taxing to consciously process everything that we see

(Beanland & Pammer, 2010). Consequently, we may often look at things without consciously being aware of what we have seen (Dretske, 2006), thus eye-tracking research can contribute to our understanding of decision-making in this context, in addition to the self-report responses in Chapter Five.

Eye-tracking technology has typically been used to assess people's eye movements when they are lying, to see whether there are changes in fixations or pupil size (Levine et al., 2006; Pak & Zhou, 2013; Schuetzler, 2012), which is informed by the emotional arousal approach of the multi-factor theory (Zuckerman et al., 1981). For example, Derrick et al. (2011) conducted a study in which some participants constructed a simulated bomb and were later shown images of arbitrary items, as well as a similar bomb to the one they had just built. They found that there were distinct differences in people's eye movements when viewing the images, which implied that eye-tracking could distinguish between those who were familiar with the bomb and those who were not. Similarly, Wang et al. (2010) found that eye tracking can be used to distinguish between honest and deceptive communication, with increases in pupil dilation when people were deceptive.

In contrast Bond (2008) used eye-tracking to assess the eye gaze behaviour of people whilst detecting deception. Two experts who had achieved accuracy rates between 80-100% in two prior instances were included in the sample. Bond (2008) found that one of the experts tended to focus on aspects of the face (lips, eyes, nose and cheek), whilst the other expert tended to focus on the arms and the top of the right leg. However, even amongst security professionals, the two experts included in the sample were unusually accurate at correctly identifying deception, therefore the extremely small sample of this portion of their study makes it difficult to extrapolate the findings to deception detection generally. Furthermore, the decision-making task in this study involved assessing videos of paroled felons who took part in a study where they lied (or were honest) about a job interview, people in their lives and their work history. Given the emphasis on gaining contextually relevant findings in this thesis, both the seated interrogation of the felons and the nature of their lies complicate the transferability of these findings to an airport context. For one, non-verbal behaviour may be more restricted when seated compared to standing/walking, potentially impacting the areas of the body which are fixated on by those assessing deception. Secondly, the experts in Bond's (2008) study had access to the felons' verbal behaviour in tangent with non-verbal behaviour, whereas this thesis aims to gain an insight of people's judgements based on non-verbal behaviour alone.

As Bond's (2008) findings show, eye tracking can provide us with more information about the subject of an individuals' attention (potentially more than merely asking them). This knowledge could be a useful tool in understanding what people look at prior to decision-making and the potential effect this may have on decision-making and accuracy. Using eye-tracking to assess decision-making may allow us to make inferences about what may have influenced accuracy and what may have led to inaccuracy, which could help to determine whether relying on non-verbal measures is suitable in airport contexts. Specifically, this study will assess people's eye gaze when identifying deception via the non-verbal behaviour of avatars when standing/walking.

7.3 Methods

7.3.1 Participants

An a priori power analysis was conducted using G*Power software (Erdfelder et al., 2009) to determine the sample size that would allow the study to have sufficient power. The significance level was set at the standard $\alpha = .05$, the power level was set to $1-\beta = .80$, and a conservative effect size was set to .25. Accounting for potential data loss due to the use of eye-tracking technology and participant withdrawal, a total sample size of 72 was determined to be sufficient. Seventy-two participants (82% women, 18% men, M_{age} = 19.65 years, SD_{age} = 1.91 years) took part in this study. Within this sample, 60% were British and 40 % were of other nationalities. All participants had normal or corrected to normal vision. Participants were recruited by responding to posters placed around the Lancaster University campus, or via the university's online recruitment system for psychological studies. All participants were required to be at least eighteen years old and have normal or corrected to normal vision. Depending on eligibility, they were either financially compensated or received course credit. Ethical approval was obtained from the Ministry of Defence Research and Ethics Committee and the Lancaster University Faculty of Science and Technology Ethics Committee.

7.3.2 Materials

Data was collected in a standard eye-tracking laboratory, using a desktop computer and the Tobii x60 eye tracker.

7.3.2.1 *Eye-tracking apparatus*

The Tobii x60 eye-tracker has a 60Hz sampling rate. This tracker was used due to the existing availability of this tracker within the laboratory. Tobii Studio software version 3.4.5 was used in conjunction with the x60 tracker for calibration and pre-analysis processing of the data.

7.3.2.2 *Eye-tracking stimuli*

The eye-tracking stimuli consisted of a 2D virtual security environment which was visually similar to those featured in Chapters Five and Six (see Chapter Three, Figure 3.3). This 2D environment was created using Unity software. Participants viewed a range of avatars (see Appendix A) whose movements were animated using the non-verbal data from participants in the study in Chapter Four. The avatar presentation was randomised for each participant. See Chapter Three, section 3.3 for further details of how the stimuli were created.

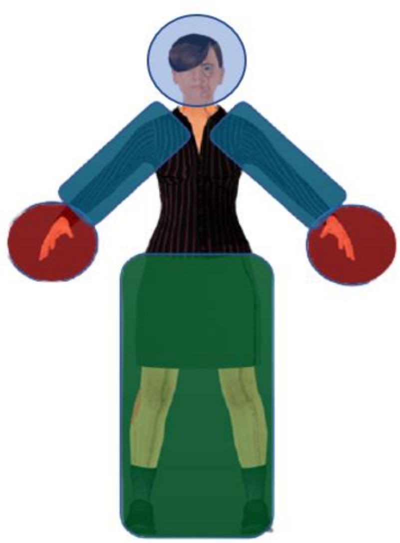
7.3.3 Measures

Participants viewed each avatar until they relayed their judgements via the desktop. As such, the following measures are proportional to the total amount of time spent observing the avatars.

7.3.3.1 *Areas of interest hits*

Six *areas of interest* were defined on each avatar within Tobii Studio. These areas were the head, right arm, right hand, left arm, left hand and legs (see Figure 7.1). The torso was not included as an area of interest due to an inability to access the laboratory to complete this part of the data processing because of the covid-19 related building closures. The leg area of interest was not separated into the left and right leg due to their closeness when the avatars were in motion, which made this difficult and redundant. The area of interest measure reflects the number of fixations on the specified part of the body.

Figure 7.1 An avatar with the six areas of interest depicted.



7.3.3.2 *Total fixations and saccades*

Tobii Studio recorded the type of eye events, either a fixation or saccade, that was present for each frame of the recording. Due to this precision, it was possible to calculate the proportion of fixations compared to saccades across each recording and relative to the deception condition. This measure accounts for the total number of fixations and saccades and is inclusive of eye movements that were not within the areas of interest in addition to those that were.

7.3.4 Design

A repeated measures design was used. The independent variables were the deception condition of the avatars (honest or deceptive) and the individual areas of interest. The dependent variable was accuracy.

7.3.5 Procedure

Informed written consent was gained before the commencement of the study. Participants were seated approximately 55cm away from the desktop during the task. Each completed a calibration trial before the study to ensure that the eye tracker was able to register their pupils in different locations of the computer screen. The participants' role and task were the same as the previous two chapters; to assess avatars approaching the security area.

A white screen with a fixation cross appeared for a few seconds before each avatar was viewed so that all participants' gaze was at the same location immediately before viewing a new avatar. Participants were instructed to focus on the fixation cross each time that it appeared on the screen. In total, they were shown a series of 24 avatars which were either honest or deceptive, in random order. Following the presentation of each avatar, the participants communicated their assessment of whether the avatar was deceptive via the desktop. After all the avatars had been viewed, a general demographics questionnaire was given and then participants were debriefed.

7.3.6 Analysis

The data was analysed in R Studio. Three separate analyses were conducted to assess: 1) Whether accuracy differed from chance- A chi-square test was used to compare the observed accuracies with chance 2) Whether the deception condition was linked with accuracy- A binomial logistic regression model was used, with deception condition as the dependent variable and accuracy as the dependent; 3) If focusing on certain areas impacted accuracy- A logistic regression model was used with deception condition as the independent variable and accuracy as the dependent variable.

7.4 Results

Overall accuracy was 49.8% and did not differ from chance $\chi^2(1, N = 72) = 0.33$, $p = .56$. There was a significant effect of deception condition on accuracy, $\beta = -0.92$, $z = -9.30$, $p <. 001$, 95% CI [-1.11,-0.73], reflecting the underlying distribution of identification of the deception (38.5%), compared to honesty (61.1%).

Accuracy was significantly worse when there were more area of interest hits on the left arm $\beta = -0.04$, $z = -2.62$, $p < .01$, 95% CI [-0.06,-0.01], but significantly better when there were more area of interest hits on the left hand $\beta = 0.03$, $z = 2.37$, $p = .02$, 95% CI [-0.01,0.06]. There was no significant effect of fixating on the other areas of interest. In light of the relationship between accuracy and the left hand and arm, further analysis compared the proportion of fixations on the left and right upper limb areas of interest and showed that the left hand was fixated on fewer times than the right hand $t(1727) = -19.22$, $p < .001$, as was the left arm compared to the right arm $t(1727) = -13.41$, $p < .001$. The leg area of interest was not linked with accuracy, but it was fixated on more than the four areas of interest on the arms, including: the left arm $t(1727) = 33.78$, $p < .001$; the left hand $t(1727) = 34.62$, $p < .001$; right arm $t(1727) = 21.76$,

$p < .001$; right hand $t(1727) = 18.10$, $p < .001$. The head was fixated on more than the legs $t(1727) = 19.02$, $p < .001$.

Table 7.1 Descriptive summary of the area of interest fixations, per body area.

Area of interest	Area of interest hits					
	Overall		Honest		Deceptive	
	M	*SD*	*M*	*SD*	*M*	*SD*
Head	24.68	21.12	26.02	21.50	23.34	20.66
Left arm	1.21	3.66	1.47	4.34	0.96	2.83
Left hand	1.42	3.84	1.56	4.12	1.27	3.54
Right arm	4.30	8.36	3.79	7.41	4.81	9.19
Right hand	5.63	8.44	5.50	8.61	5.77	8.28
Legs	12.61	13.34	11.84	12.92	13.71	13.71

Note. Area of interest hits are presented overall and relative to the deception condition

7.5 Discussion

The results of this study showed that the overall accuracy of identifying deception was poor. The deceptive avatars were linked with a greater incidence of fixations on the left arm, right arm and legs. Accuracy varied depending on the area of interest that was fixated on, with reduced accuracy when fixating on the left arm whilst fixating on the left hand had a positive effect on accuracy. Participants looked at the legs significantly more than the other areas of interest, although this was not linked with better accuracy.

It is important to note that the left arm and left hand did not differentiate truth-tellers from deceptive people in Chapter Four. Movement of the left arm and hand merely distinguished between two of the deceptive conditions, package and weapon. There were contrasting effects on accuracy between looking at the left arm compared to looking at the left hand, and although Chapter Four did not find that the arms or hands distinguished truth-tellers from liars, previous research has found that people believe that finger and hand movements increase when people are deceptive (Vrij & Semin, 1996). Likewise, upper limb movements were cited in the self-report as cues that were used to inform decision-making in Chapters Five, which may explain why these areas were fixated on. It is important to stress that there is some difficulty with the interpretation of fixations because the results do not explain why the effect on the accuracy of looking at the left hand did not apply to the arm.

The question remains of why accuracy increased when looking at the left hand and not the right, especially given that there were significantly fewer fixations on the left hand compared to the right. There are conflicting findings on gaze preference, with some studies finding a preference for looking at the left visual field (Durgin et al., 2008; Guo et al., 2009) whilst others suggest a preference for the right (Gorbunova & Falikman, 2019). This study's findings support the notion of a preference for the left, given the higher number of fixations for the right arm and hand, which appeared on the left visual field of the screen. The majority of people are right-handed (Holder, 2001), which may cause people to focus more on the right hand of others. Given the somewhat sporadic nature of the effect of looking at the hand compared to the arm on the accuracy, these findings should be interpreted cautiously, as this may be a statistical anomaly.

Much like previous chapters, another potential area for research could include a more occupationally diverse sample. This could look at whether eye movements vary between security personnel and laypeople, if this interacts with the deception condition and whether this has an impact on accuracy. Given that security personnel should have greater experience of assessing people for the presence of deception than the layperson sample in this chapter, different eye gaze patterns may have emerged had they been included in this study. However, due to findings that some personnel in the security sector have misconceptions about cues of deception (Denault et al., 2020; Strömwall & Granhag, 2003; Vrij & Semin, 1996), their gaze behaviour may show fixations that are not linked with deception. Nevertheless, such research into the gaze of professionals is an often-overlooked step in working towards our understanding of detection accuracy and may prove useful.

7.5.1 Limitations

The use of the virtual reality avatars and security checkpoint stimuli was convenient due to the ease of accessibility and incorporation of this technology within this thesis. Whilst it provides a stepping stone in understanding the eye gaze behaviour of people observing potential deception, it poses a limitation in terms of knowing whether the same eye movement behaviour would be present when viewing non-virtual people. Due to the limited range of studies comparing the outcomes of virtual reality studies with the real world, it is difficult to judge whether these findings are wholly applicable to viewing real people. Accordingly, a study measuring eye movements when viewing CCTV footage or recordings of people in a laboratory setting would be a useful step in furthering the utility of this research.

7.5.2 Summary

In conclusion, eye tracking research can provide some useful information in an area of deception research that has rarely been investigated, the gaze behaviour of the observers. This study shows the importance of gaining a more objective insight into what may be influencing people's decision-making. The results highlight a potential issue of fixating on specific parts of the body because the way that perceived non-verbal movements are processed and interpreted may not affect accuracy at all or positively. This may be due to the ambiguity with deciphering others' behaviour and in tangent with the reliability/validity of non-verbal cues, may explain some of the poor results of deception detection in the literature.

Part V: Chapter Eight- General Discussion

The overall findings of this thesis suggest that the assumption that judgements based on non-verbal behaviour are a reasonable method of determining if others pose a threat is flawed, specifically in the absence of valid non-verbal markers. Using the framework of the lens model to structure this research (Brunswik, 1952), the non-verbal measures which were assessed were not valid markers of deception and people's accuracy was poor and reliant on superficial factors or irrelevant factors. This final chapter will provide a brief overview of the motivation (from the literature and airport practice) for this thesis; evaluate the use of virtual reality and automated measurement tools; summarise the results and theoretical implications from Chapters Four to Seven; consider the practical implications of the findings and elaborate on the overall limitations of this research. Suggestions of some avenues for future work in this field will be mentioned throughout.

8.1 Overview

Within many contexts, there is a need to be able to distinguish between those who are telling the truth and those who are lying. Though the deception field has sought to identify behavioural markers that can distinguish between honesty and deception, they tend to be unreliable or limited in their diagnostic value (Nortje & Tredoux, 2019; Vrij et al., 2019). Regardless, people tend to rely heavily on non-verbal behaviour to make assessments about other's intentions and whether they pose a threat (Bogaard et al., 2016; Johnson, 2007). However, the literature on deception detection suggests that people's ability to detect deception is poor (Bond & DePaulo, 2006; Vrij, 2008a).

Within some airport settings, a reliance on non-verbal cues may be promoted by security protocols that focus on non-verbal behavioural abnormalities to justify approaching people for questioning (U. S. Government Accountability Office, 2010). In practice, there are widespread accusations of prejudicial decision-making, influenced by superficial factors such as race, ethnicity and religion (ACLU, 2017; Meyer, 2010). Along with the moral issues with this type of discrimination, this may also detrimentally influence the objectivity of decision-making. Considering these issues, the use of non-verbal measures may be inappropriate in an airport context given the potential for false positives (falsely identifying someone as potentially posing a threat) and false negatives (failing to identify someone who poses a threat).

Airports were the specific environment of interest in this thesis. There was an issue with blindly applying the findings of some of the research in the field of deception to an airport security checkpoint context. A lot of our knowledge of non-verbal cues is derived from laboratory research which may lack direct contextual relevance when trying to understand deception and deception detection behaviour in an airport environment as opposed to the seated police interrogation methodology that is common in the literature. To gain a clearer understanding of the suitability of non-verbal behaviour to an airport environment, virtual reality was used as a method of contextual priming to attempt to evoke real-world behaviours. This thesis aimed to assess whether non-verbal behaviour differed in a virtual airport environment (Chapter Four); peoples' ability to identify deception in others and the factors that influenced decision-making (Chapters Five, Six and Seven).

8.2 The use of technology in this research

8.2.1 Virtual reality

As discussed in Chapter Two, a lot of the work on deception has revolved around a seated police interrogation paradigm, within a laboratory environment. The applicability of some findings to our understanding of non-verbal behaviour in airport environments is difficult because of the potential impact of contextual differences. To focus this research on an airport environment, virtual reality was used to reduce the salience of the laboratory environment and encourage contextual priming to evoke realistic behaviour (see Chapters Two and Three). The use of virtual reality in this thesis was influenced by the increasing importance of conducting contextually relevant research, particularly within the deception field (Vrij et al., 2019), especially because of the need to contribute towards ensuring that airport security practices are valid and based on sound research.

Virtual reality has the potential to bridge the gap between more traditional laboratory environments and field settings (de Groot et al., 2020). Inevitably, laboratory-based security studies cannot capture all the situational variables that are present in the real world (Sotirakopoulos et al., 2011). One of the issues within deception research is that it is difficult to generalise specific measures of deception to other contexts, which may influence the reluctance of practitioners to trust the findings of laboratory studies (Buckley, 2012). Replicating environments that may otherwise be difficult to gain access to (Brookes et al., 2019; Rebelo et al., 2012) is relatively easy within virtual reality. Also, the ease of replicating the visual, auditory and some tactile stimuli allowed for the simulation of an airport in a laboratory that otherwise would not have been possible within the confines of the available space during this research period. In addition, compared to field experiments, using virtual reality has the advantage of enabling researchers to maintain control of the variables to an extent that may not be possible in an airport because of the multitude of variables constantly at play (Pan & Hamilton, 2018).

Although using virtual reality may address some of the aforementioned issues in more traditional studies, it would be remiss to neglect the discussion of some of the limitations of virtual reality. While virtual reality in some instances may be an improvement in terms of the contextual relevance of experimental control and reproducibility compared to other methods (Pan & Hamilton, 2018), few studies have assessed the comparability of virtual reality studies to the real world. Difficulties securing access to an airport prevented the inclusion of a comparison of behaviour in virtual reality with a field study within this thesis, so one must be cautious with generalising these findings to real security checkpoints without further research to support doing so. Previous research has suggested that in some contexts people behave similarly in real life and virtual reality (Armougum et al., 2019; Bhagavathula et al., 2018; Xu et al., 2021), but the use of virtual reality and comparison studies with real-world behaviour is still fairly novel within the deception domain. This is an important avenue of future research to validate the comparisons of deception-related behaviours in virtual reality and the real world. With this thesis, it is difficult to ascertain whether judging virtual avatars elicits the same responses as judging real people. Although virtual reality was used to reduce the salience of the laboratory, ultimately within deception research the participants are aware that they are taking part in a study, as is the case with most types of human research. Anecdotally, some people did report forgetting that they were in the laboratory but there is the potential that people's behaviour was still not the same as it would have been in a real airport.

The virtual reality immersion scores were poorer in Chapters Five and Six when people judged whether the avatars were honest or deceptive, compared to Chapter Four. The reduction in immersion was likely due to the reduced interaction in these virtual environments compared to Chapter Four which had more of a mixed reality environment (Pollard et al., 2020) due to the incorporation of tactile stimuli, walking in virtual reality and verbal interaction with other avatars. The reduced immersion may have impacted the extent to which the accuracy rates and the effect of the information reduction techniques were reflective of real life. However, the accuracy rates and the truth bias underpinning them were consistent with the literature, so immersion in Chapters Five and Six is unlikely to have had a detrimental impact on the conclusions from these chapters. Overall, virtual reality was a helpful addition to this research and may be beneficial to other aspects of deception research, but the conclusions must be approached cautiously in relation to the real world. Future use should aim to increase immersion by introducing more interactive elements in their designs (Pan & Hamilton, 2018; Slater et al., 1994) and include real-world comparison conditions.

8.2.2 Measuring non-verbal behaviour

Non-verbal deception research has typically measured changes in behaviour using manual methods of coding and processing the data (Klaver et al., 2007), in which people observe videos and manually code for specific changes in movement. This thesis used automated methods of measuring non-verbal behaviour via the use of Xsens motion-capture suits, and an automated method of processing the data output from the Xsens suits using the AMAB method (Poppe et al., 2014). Reducing the potential for subjective researcher influence to affect the non-verbal data, either consciously or subconsciously (Luke, 2019), was one incentive for avoiding manual coding methods. Using motion-capture and the AMAB method was beneficial because it allowed for small changes in behaviour to be assessed, which would not have been possible with the naked eye. Furthermore, the AMAB method allows for a more gestalt approach to assessing differences in non-verbal behaviour compared to manual observation methods. AMAB can compute changes in one area of the body relative to another, taking into account that joints seldom move independently, thus providing an overall picture of non-verbal behaviour as opposed to homing in on more prominent attributes.

The Xsens suit consists of a lycra zipped t-shirt and several sensors placed on self-fastening straps worn around the limbs. Though anecdotally, some participants reported forgetting that they were wearing it, it is possible that not everyone perceived the suit to be unintrusive. It could be reasoned that an unintended consequence of the suit is that it could impact some people's non-verbal behaviour. Much like the use of virtual reality, future work should aim to measure the similarity between movement when wearing the suit and not.

8.3 Empirical summaries and theoretical implications

8.3.1 Deception cues

There is an assumption within the deception literature that non-verbal behaviour should differ between deceptive and honest people because of cognitive load, emotional arousal and/or behavioural control (Zuckerman et al., 1981). Chapter Four aimed to assess whether that assumption would apply to a virtual airport environment, or whether the contextual pressure of an airport would render other models more insightful, such as the self-presentation theory (DePaulo, 1992).

8.3.1.1 *Non-verbal behaviour*

The first objective was to investigate whether there were valid cues for detecting deception. This question was answered in Chapter Four in an experiment that used both the Xsens motion capture system and virtual reality to assess whether there were changes to non-verbal behaviour and response latency (a conceptual replication of a previous study) when people were deceptive. This study had an honest condition and three deceptive conditions: future action, package and weapon. There were no distinctions between the honest and deceptive conditions in terms of the non-verbal measures: movement per minute, cadence, step length, speed, centre of mass displacement or segment displacement. Segment displacement only revealed differences between the package and weapon condition, specifically with more movement of the left forearm and hand in the weapon condition. Response latency was shorter in the honest condition, only compared to the future action condition. The response latency finding contradicts a conceptually similar virtual reality study where people had shorter response latencies when they were deceptive (Mapala et al., 2017).

The findings from Chapter Four potentially highlight the importance of environmental context and the subsequent types of deception when researching with the aim of generalising to a particular environment. The null results of non-verbal markers contrast with other studies that have found significant differences in a range of non-verbal measures, also using motion-capture (Poppe et al., 2014). Eapen et al., 2010 found that non-verbal behaviour differed when deceptive, but this task was about participants lying about their performance on a maths test and having witnessed an accident. Similarly, Matsumoto et al. (2015) found that non-verbal measures distinguished between people that were deceptive and honest. It is possible that these different results were due to the different methodologies, virtual reality compared to a physical security checkpoint in a laboratory, or task differences (Eapen et al., 2010). Therefore, laboratory salience may have contributed to the discrepancy with this thesis. Likewise, Poppe et al., 2014 found that joint displacement was different between truth-tellers and liars, with their methodology being directly relevant to interrogation proceedings. Though previous research may be relevant to the aims and contextual parameters set by the researchers, this thesis aimed to address the validity of assessing non-verbal behaviour in airport environments. Hence, while acknowledging the bleak picture painted by this thesis about the assumption that non-verbal behaviour would differ in an airport context, further research is needed to replicate and understand the discrepancy between these findings and other literature.

Although the null non-verbal findings do not support the multi-model theory (Zuckerman et al., 1981) because the cognitive load, emotional arousal and behavioural control presumably did not lead to behavioural distinctions in honest and deceptive people, it is important to note that none of these three factors were measured directly in thesis research. Consequently, one interpretation of the null findings is that these three factors did increase for participants in the deceptive conditions, but that any increase simply did not have a significant effect on their non-verbal behaviour. Another interpretation of the findings is that the task did not increase cognitive effort, emotional arousal or behavioural control at all, or not to a significant degree in the deceptive conditions, compared to the honest condition. Within deception research, as we strive to design more realistic studies, it is equally important that studies do not artificially increase the propensity for increased cognitive effort, emotional arousal or behavioural control which goes beyond the extent that those factors are experienced in real-world contexts for the sake of experimental manipulation. Consequently, one of the theoretical implications of this research is that it may cast doubt on the extent to which some of our theoretical assumptions may be reflective of laboratory-induced behaviours and not the real world. However, caution should be applied with this theoretical implication based on one piece of research.

The null non-verbal findings could lean toward one aspect of the self-presentational theory (DePaulo, 1992), which emphasises that there may be more similarities than differences between liars and truth-tellers. It is difficult to ascertain the cause of the non-significant difference between the honest and deceptive conditions though. It could reflect both the honest and deceptive participants modifying their behaviour due to the airport context (Blackwood et al., 2015). However, this cannot be definitively concluded because there were no baseline measures of the participants' non-verbal behaviour. The between measure design of Chapter Four limits the conclusions that can be drawn from the findings. Particularly in terms of the self-presentational theory, whether the null results are reflective of all the participants modifying their behaviour to present a better image, or whether all participants were equally unaffected by the scenario, may be an interesting follow-up. Future research may wish to incorporate a within-measures design to compare each individuals' baseline behaviour to the post-experimental intervention behaviour to form a clearer assessment of the mechanisms underlying non-verbal behaviour in this context.

8.3.1.2 *Response latency*

The longer response latencies when people were deceptive partially aligns with previous findings (Debey et al., 2015; Sporer & Schwandt, 2007; Suchotzki et al., 2013; Walczyk et al., 2013). There is only partial support for the cognitive load approach regarding response latencies because they were longer only in the future action condition compared to the honest condition. Contrasting with the null results in the non-verbal measures this could suggest that the multi-model may be dependent on both context and whether verbal or non-verbal measures are used. The difference in response latency between the honest and future action condition may have been the result of the increased cognitive load caused by having to suppress the truth in their condition, that they were intending to meet with someone to commit an illegal act, and maintain the deceptive narrative (Zuckerman et al., 1981). Lying about future intentions may have resulted in a distinguishable response latency because there is a degree of uncertainty over anticipated future actions compared to the more concrete reality of having a prohibited item on one's person. Due to this degree of uncertainty over the events of an anticipated future illegal act, people may be more cautious when responding. The null latency result of the other deceptive conditions is interesting as it could suggest that response latency may not be able to distinguish between honest people and those who may be carrying prohibited items, such as a weapon or suspicious package. However, again this null result should be interpreted cautiously.

Much of the research on response latencies have focused on deception concerning current or past actions (Geven et al., 2020; Sheridan & Flowers, 2010), though research on deception about future intent typically revolves around the content of verbal communication (Kleinberg et al., 2017), not response latency. These findings suggest that the effect of deception on response latency may be mitigated by the type of lie, with different effects due to the concealment of an object compared to misleading information about a future intention. Thus, the assumptions about response latency (Suchotzki et al., 2017) may not be generalisable to all forms of deception. However, a similar virtual reality study in which people lied about possessing a prohibited item found that there were differences in response latency between the honest and deceptive condition, and the effect of deception on response latency was in the opposite direction (Mapala et al., 2017). Mullin et al. (2014) also found distinguishable differences in response latency when people lied about possessing a prohibited item, suggesting that the effect of deception on response latency is not limited to lies about future events as the findings from this thesis may suggest. Further replication would be helpful before speculating about why the response latencies of participants in the weapon and package condition were unaffected by deception.

8.3.2 Accuracy identifying deception and the factors that influence decision-making

Accuracy identifying deception was measured in Chapters Five, Six and Seven. Despite the lack of significant non-verbal measures distinguishing honesty from deception in Chapter Four, the second objective addressed in Chapter Five sought to assess how accurate people were at identifying deception based on non-verbal behaviour and which factors they reported using to influence their judgements. There are many facets of non-verbal behaviour, all of which could not feasibly be assessed within this project. The accuracy measure in Chapters Five investigated whether accuracy in the absence of the validity of the non-verbal measures from Chapter Four was no different from guessing, or whether people were good at identifying deception, hinting at an inherent ability to identify deception based on something else. Indirectly, Chapter Five to Seven aimed to contribute towards having a greater understanding of whether non-verbal judgements via human observation are valid in an airport setting, given the unreliability of the diagnostic value of non-verbal behaviour and deception (shown in Chapter Four).

The overall accuracy findings from Chapters Five to Seven agreed with the literature (Bond & DePaulo, 2006) that accuracy was poor and either did not differ from chance levels or was worse than chance. The overall implication of these findings is that people are poor at detecting deception in instances when they must rely on non-verbal behaviour. The underlying distribution of the poor accuracy did show support for the truth-default theory (Levine, 2014b).

The non-significant difference in overall accuracy from chance could allude to the fickleness of identifying deception from non-verbal behaviour. The underlying distribution of the poor accuracy rate supports the truth-default theory (Levine, 2014b) likely because most populations, other than law enforcement (Masip & Herrero, 2017), generally assume that people are honest more often than they lie. The accuracy findings extend our knowledge of the poor suitability of assessments based on non-verbal behaviour, to a simulated airport environment. It would stand to reason that if accuracy above chance levels was not evident in the noise-free conditions of a virtual laboratory experiment, then it is unlikely that judgements based on non-verbal behaviour would be any more suitable in an actual airport.

The thematic analysis found that six themes were used to influence decision-making: physical appearance; disposition; walking behaviour; body positioning; looking behaviour and upper limb movement. The responses from the thematic analysis revealed often conflicting perceptions of whether one of the themes, for example, disposition, implied deception. The discrepancies in perception may be attributed to individual differences in the participants and are somewhat consistent with some of the critiques of the SPOT program, namely that an innocuous trait or behaviour can be erroneously deemed to be indicative of guilt, due to the subjective nature of detecting threats from non-verbal behaviour. The contrasting responses from the thematic analysis convey the difficulty of objectively assessing non-verbal behaviour and the ease with which perceptual differences, potentially mitigated by one's own experiences, can impact whether behaviours are interpreted as being indicative of deception. On the other hand, due to the predominantly student population that the thematic responses were garnered from, it is imaginable that security personnel may have had different interpretations of the non-verbal behaviours. Nonetheless, whilst the content of the responses may vary both on an individual basis and reflective of occupation (akin to the lie bias (Masip & Herrero, 2017)), the contrast in the interpretation of factors such as disposition reflects an overall important hindrance of relying on people to detect deception through non-verbal means. Future work may wish to gain a qualitative insight into the way that security personnel interpret the behaviour of others.

8.3.3 The effect of reducing subjective influence on decision accuracy

The third objective, to improve accuracy was addressed in Chapter Six and was influenced by the poor accuracy in Chapter Five, which was no different to guessing (chance) and the thematic analysis which suggested a reliance on superficial factors. Chapter Six investigated whether the poor accuracy rate found in Chapter Five and the wider literature could be improved, relative to chance. The incorporation of two techniques of information reduction: reduced superficial information; and imposed time constraints were used. The effectiveness of these two techniques which aimed to reduce the salience of distracting information was assessed by comparing the accuracy rates to chance levels. Overall, accuracy was no different from chance levels for the reduced superficial information condition and worse than chance levels for the imposed time constraints condition.

Chapter Six contradicted Hartwig and Bond's (2011) suggestion that measures to increase the reliance on intuition, such as by reducing the presence of non-pertinent information, may boost accuracy. Instead, Chapter Six found that placing time constraints on exposure to the visual stimuli did not improve accuracy relative to chance. This outcome may have been due to the method of inducing unconscious processing rather than a flawed assumption that unconscious processing would increase accuracy. It is possible that the time constraints were too extreme to positively affect accuracy. However, although Reinhard et al. (2013) and Albrechtsen et al. (2009) found that inducing unconscious decision-making by manipulating the time for deliberation or stimuli exposure improved the accuracy of detecting deception, other research has also failed to replicate such findings (Moi & Shanks, 2015; Wu et al., 2019), perhaps suggesting that this intervention is not beneficial in this context and/or using this methodology.

Chapter Six also tried to boost deception detection accuracy by using point-light displays to reduce the salience of irrelevant visual information. This intervention also failed to boost accuracy, which conflicts with previous findings that reducing superficial information increased accuracy (Park et al., 2019) and point-light display studies where accuracy rates were at least 70% (Runeson & Frykholm, 1983). The difference between the type of tasks may explain this lack of a significant finding. For example, whilst the visual stimuli in Runeson & Frykholm 's (1983) study involved carrying a box, the visual stimuli in Chapter Six were simply of people walking, as one would at a security checkpoint. Both involve deception, but the box task is a measure of deception due to a misleading physical action, whereas the deception in this thesis' point-light display condition was due to the concealment of a prohibited object or future intention. Hence, the use of the point-light displays may not benefit accuracy when detecting deception that does not involve deliberate manipulations of the body.

8.3.4 Eye gaze behaviour

The fourth and final objective was to assess what people looked at prior to decision-making and how this relates to accuracy. This was assessed in Chapter Seven using eye-tracking equipment to gain a more objective insight of assessing what people use to inform their decision-making, as opposed to the self-report measures in Chapter Five. The results found that people's fixations on the left arm and left hand were linked with accuracy, despite joint displacement findings in Chapter Four not supporting the relevance of either area as distinguishing between honesty and deception. The effect of these fixations on accuracy differed. People were more accurate at detecting deception when they fixated on the left hand, whereas accuracy was worse when fixating on the left arm. These results were despite participants looking at the right hand and arm more than the left hand and arm and despite more fixations on the legs compared to all other areas of the body (except the head). Overall accuracy was again no different from chance, at 49.8%.

The eye-tracking findings of fixations on the hands and arms and the relationship with accuracy are not justified in this thesis since these areas were not indicated as distinguishing between truth-tellers and liars in Chapter Four, though they did distinguish between the weapon and package condition. However, these findings support previous literature which suggests that deception influences hand movements (Sartori et al., 2016) and commonly held beliefs that hand movements are a cue of deception (Strömwall et al., 2004), so the fixations may be reflective of beliefs about deception markers, specifically related to upper limb movement. The legs were fixated on the most compared to the other areas of the body, although this was not linked with accuracy. Since none of the measures relating to the legs in Chapter Four was significant, the legs being the primary area of focus does not support Hartwig and Bond's (2011) idea that people look at valid cues and poor accuracy reflects being unable to consciously make decisions based on valid markers. It is important to note that this type of eye-tracking research shows what was fixated on before making a judgement, but it cannot be assumed that these fixations directly impacted decision-making.

8.3.5 Implications summary

The abovementioned findings have theoretical implications regarding the mechanisms underpinning non-verbal behaviour when deceptive. In Chapter Two, two main theories of contention were discussed, the multi-factor theory (Zuckerman et al., 1981) and the self-presentational theory (DePaulo, 1992). Though the design limitations of Chapter Four prevent explicit support for the self-presentational theory, the theoretical implications of this thesis support previous notions of the idiosyncratic nature of non-verbal behaviour (Jupe & Keatley, 2019; Porter & Ten Brinke, 2010) and thus the complexity of any individual theory providing a sufficiently conclusive explanation of non-verbal behaviour in a range of contexts (Vrij et al., 2019). Furthermore, this thesis provides support to the findings that people are poor at identifying deception and that methods of inducing intuitive decision-making are not effective (Moi & Shanks, 2015) at producing accuracy rates that can justify protocols reliant on non-verbal behaviour. Given that in airport contexts non-verbal behaviour can be more readily available, hence a reliance on it, Chapter Six sought to see whether even in the absence of valid cues, information reduction techniques could boost accuracy above chance levels. The rationale behind this was partly based on the common assertions that people are intuitively good at determining whether people are deceptive (Hartwig & Bond, 2011; Masip & Herrero, 2017; Pinizzotto et al., 2004; Reinhard et al., 2013). However, such judgements are often deemed to be reflective of prejudice based on racial, religious or ethnic categorisations (ACLU, 2017; Meyer, 2010). The lack of improvement in accuracy when reducing the reliance on biasing heuristics to allow for a more intuitive assessment of deception (Albrechtsen et al., 2009; Hartwig & Bond, 2011; Reinhard et al., 2013) leads to the conclusion that some of the theories regarding the benefit of intuition are not suited for application in a high-security context.

8.4 Practical implications

In high volume areas, often non-verbal behaviour is the first method of assessment before security personnel decide that a person may need to be approached or require further investigation to determine their intentions (U.S. Government Accountability Office, 2010). The limited scientific underpinning for behavioural detection methods used at some security checkpoints and within other legal contexts (Jupe & Denault, 2019; U.S. Government Accountability Office, 2010) is a concern. The invalid protocols as well as the false beliefs about cues of deception by those within the police, prosecutorial and judiciary industries (Jupe & Denault, 2019; Strömwall & Granhag, 2003) points to a larger issue with which academia could potentially help.

Although other studies of non-verbal behaviour have found valid cues, the context of these studies may have contributed to their findings and they may not apply to security checkpoint environments. The knowledge gained from research is not being successfully disseminated beyond some academic circles or is inappropriately interpreted (Jupe & Denault, 2019). This notion is evident due to the resistance to address the criticisms of the SPOT program

(U.S. Government Accountability Office, 2013) and the lack of knowledge held by those detecting deception in security settings (Denault et al., 2020). The reliance on invalid cues in practice may be due to difficulty understanding the applicability of laboratory findings to real security environments and doubts from those within the security industry about the usefulness of laboratory-based findings (Buckley, 2012). The use of virtual reality within this thesis aimed to increase the contextual relevance of the research and therefore reduce concerns about the irrelevance of laboratory-based deception studies to the security industry. Though Hartwig and Bond's (2014) meta-analysis suggests that such concerns are overstated by those in security, this research used virtual reality to replicate a security checkpoint to try to quell these. It was intended that the transferability of these findings to security contexts may be more apparent and contribute towards lessening the reliance on pseudoscience in the real world, though real-world comparisons and a more occupationally diverse sample are needed to justify this.

The null non-verbal results may have important practical implications, as it shows that non-verbal differences may not always be found, nor are they advantageous to decision-making, casting doubt on their inclusion in security protocols. If an automated measure which assesses behaviour at the minute level could not find significant differences in non-verbal behaviour, how can we expect security personnel to make judgements based on naked-eye observations of non-verbal behaviour? A poor ability to detect deception means that honest people may be incorrectly assumed to be deceptive. These errors can have long-term consequences, as honest people can be indefinitely detained (Hickman, 2005; Proulx, 2005; Stewart, 2005) or wrongfully convicted (Innocence Project, 2019) based on initial false assumptions which spiral out of control. On the other hand, poor detection abilities mean that deception may go undetected when people do pose a significant threat. Questionable reliability/validity of non-verbal diagnostic cues and the poor ability to detect deception using measures to reduce any noise around detection based on non-verbal behaviour (Chapter Six) could mean that people's reliance on non-verbal behaviour may be a hindrance to their ability to identify deception.

8.5 Ethical implications

One of the main ethical implications of deception research, particularly that which investigates future intent as in Chapter Four, regards the potential abuse of applications in the real world. For example, with the longer response latencies in the future action condition in Chapter Four, a crime has not been committed at the point in time where the discrepancies in response latency are evident. Questions emerge in terms of the acceptable follow-up actions regarding the use of response latency measures that aim to identify deception about criminal future intent. Would such a measure be used to permit further surveillance? A future intent measure such as this could potentially be susceptible to abuses of power, suffering from the same critique that plagues the 'preventative' stop and search/frisk measures particularly in London and New York (Bradford & Tiratelli, 2019) over the dubious justification for further intervention.

8.6 Overall limitations

The limitations of this work have been mentioned throughout the empirical chapters and within this chapter, but this section will elaborate on the overarching limitations of this thesis. A debate around the representativeness of the participant pool of a lot of behavioural experiments may apply to this thesis. Researchers have argued that most people who volunteer for psychological experiments are western, educated, industrialised, rich and from democratic nations (WEIRD; Dan, 2010). The WEIRD-ness of participants is argued to limit the extent to which the results of studies can be generalised to the diverse groups of people that are present at airports. Although not formally recorded, a substantial amount of the participants in this research were not from western nations, so culturally the results may be generalisable to an extent. However, most participants were undergraduate or postgraduate students, meaning that they were educated and possibly were from industrialised nations, thus complicating generalisations beyond WEIRD populations. Other aspects of the participants may limit the generalisability of these findings.

The age range of the sample population in comparison to the more diverse range of people at security checkpoints may limit the scope of generalisability. Though the samples used were culturally and ethnically diverse, most participants were students under the age of 30. As mentioned previously, there are non-pathological declines in cognitive ability as people age, which can start from as early as 30 years old (Deary et al., 2009). Hence, age may have an impact on the effect of engaging in deception on the non-verbal and verbal behaviour that people exhibit, as well as the ability to identify deception. Therefore, the generalisability of these findings to people over 30 may be negatively impacted.

The sole use of a layperson participant pool poses a potential limitation on the generalisability of this research. The lack of security personnel in the samples in this thesis may impact the conclusions that can be drawn. Although some research has found that accuracy of detecting deception can vary in security personnel depending on the extent of their training and the length of their experience (Ekman & O'Sullivan, 1999), other findings suggest that very small populations have accuracy exceeding the abilities of a predominant student population (Bond, 2008), to the extent that security personnel are not much better than lay-populations. Consequently, the student population should not detrimentally affect the extent to which these results apply to other populations.

8.7 Concluding remarks

This thesis has shown that neither non-verbal behaviour nor response latency may be able to definitively distinguish truth-tellers from liars. Peoples' ability to identify deception in others was poor, influenced by subjective perceptions that did not improve with information-reduction techniques, and irrelevant areas of the body were fixated on before decision-making. Virtual reality was used to impact the contextual relevance of this research and motion capture technology was chosen to gain precise measurements of the non-verbal cues. The null findings regarding the diagnostic value of the non-verbal measures assessed and the poor accuracy reported throughout this thesis casts doubt on the suitability of non-verbal measures deception detection at security checkpoints, as is used in the TSA's SPOT program. The culmination of these findings leads to the conclusion that relying on non-verbal behaviour to identify deceit may be an unjustifiable method.

References

Abernethy, B., Jackson, R. C., & Wang, C. (2010). The perception of deception: The role of kinematic and other information in detecting deceptive intent within movements. *Journal of Sport & Exercise Psychology*, *32*, 56.

ACLU (2017). Bad trip debunking the tsa's "behavior detection" program. In *American Civil Liberties Union*. https://www.aclu.org/report/bad-trip-debunking-tsas-behavior-detection-program

Al-Amri, M., Nicholas, K., Button, K., Sparkes, V., Sheeran, L., & Davies, J. L. (2018). Inertial measurement units for clinical movement analysis: Reliability and concurrent validity. *Sensors*, *18*(3), 1–29. doi.org/10.3390/s18030719

Albrechtsen, J. S., Meissner, C. A., & Susa, K. J. (2009). Can intuition improve deception detection performance? *Journal of Experimental Social Psychology*, *45*(4), 1052–1055. doi.org/10.1016/j.jesp.2009.05.017

Allison, A. L., Peres, J. C., Boettger, C., Leonbacher, U., Hastings, P. D., & Shirtcliff, E. A. (2012). Fight, flight, or fall: Autonomic nervous system reactivity during skydiving. *Personality and Individual Differences*, *53*(3), 218–223. doi.org/10.1016/j.paid.2012.03.019

Alshak, M. N., & Das, J. M. (2020). *Neuroanatomy, sympathetic nervous system*. Statpearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK542195/

Ambady, N. (2010). The perils of pondering: intuition and thin slice judgments. *Psychological Inquiry*, *21*(4), 271–278. doi.org/10.1080/1047840X.2010.524882

American Polygraph Association (1997). The validity and reliability of polygraph testing. In *Polygraph*, *26*(4), 215-239.

American Psychological Association. (2019). *Stress effects on the body: Nervous System*.

Ayada, C., Toru, U & Korkut, Y. (2015). The relationship of stress and blood pressure effectors. In *Hippokratia*,*19*(2), 99-108.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. doi.org/10.1016/S0079-7421(08)60452-1

Bailenson, J. N., Davies, A., Blascovich, J., Beall, A. C., Mccall, C., & Guadagno, R. E. (2008). The effects of witness viewpoint distance, angle, and choice on eyewitness accuracy in police lineups conducted in immersive virtual environments. *Presence*, *17*(3), 242–255.

Baker, E. (2002). Flying while arab- racial profiling and air travel security. *Journal of Air Law and Commerce*, *67*, 1375–1406.

Baldwin, J. (1993). Police interview techniques: Establishing truth or proof ? *The British Journal of Criminology*, *33*(3), 325–352.

Bartlett, M. S., Littlewort, G. C., Frank, M. G., & Lee, K. (2014). Automatic decoding of facial movements reveals deceptive pain expressions. *Current Biology*, *24*(7), 738–743. doi.org/10.1016/j.cub.2014.02.009.

Beanland, V., & Pammer, K. (2010). Looking without seeing or seeing without looking? Eye movements in sustained inattentional blindness. *Vision Research*, *50*(10), 977–988. doi.org/10.1016/j.visres.2010.02.024

Benner, L. A. (2009). The presumption of guilt: systemic factors that contribute to ineffective assistance of counsel in california. *California Western Law Review*,45(2), 1-111. https://scholarlycommons.law.cwsl.edu/fscontentdownloaded/printedfromheinonline

Bhagavathula, R., Williams, B., Owens, J., & Gibbons, R. (2018). The reality of virtual reality: A comparison of pedestrian behavior in real and virtual environments. *Proceedings of the Human Factors and Ergonomics Society*, *3*, 2056–2060.

doi.org/10.1177/1541931218621464

Bideau, B., Kulpa, R., Vignais, N., Brault, S., Multon, F., & Craig, C. (2010). Using virtual reality to analyze sports performance. *IEEE Computer Graphics and Applications*, *30*(2), 14–21. doi.org/10.1109/MCG.2009.134

Blackwood, L., Hopkins, N., & Reicher, S. D. (2015). ‘Flying while muslim’: Citizenship and misrecognition in the airport. *Journal of Social and Political Psychology*, *3*(2), 148–170. doi.org/10.5964/jspp.v3i2.375

Blair, J. P., Levine, T. R., & Shaw, A. S. (2010). Content in context improves deception detection accuracy. *Human Communication Research*, *36*(3), 423–442. doi.org/10.1111/j.1468-2958.2010.01382.x

Blalock, G., Kadiyali, V., & Simon, D. H. (2007). The impact of post-9 / 11 airport security measures on the demand for air travel. *Journal of Law and Economics*, *50*(4), 731–755.

Bogaard, G., Meijer, E. H., Vrij, A., & Merckelbach, H. (2016). Strong, but wrong: Lay people’s and police officers’ beliefs about verbal and nonverbal cues to deception. *Plos ONE*, *11*(6), 1–19. doi.org/10.1371/journal.pone.0156615

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*(3), 214–234. doi.org/10.1207/s15327957pspr1003_2

Bond, G. D. (2008). Deception detection expertise. *Law and Human Behavior*, *32*(4), 339–351. https://doi.org/10.1007/s10979-007-9110-z

Bouchrika, I., Goffredo, M., Carter, J., & Nixon, M. (2011). On using gait in forensic biometrics. *Journal of Forensic Sciences*, *56*(4), 882–889. doi.org/10.1111/j.1556-4029.2011.01793.x

Bradford, B., & Tiratelli, M. (2019). Does stop and search reduce crime? In *UK Justice Policy Review Focus*. https://www.crimeandjustice.org.uk/sites/crimeandjustice.org.uk/files/Does stop and search reduce crime.pdf

Bradley, S. (2014). *Bottom-up visual perception: what saccadic eye movements suggest for designers.*Vanseo Design. https://vanseodesign.com/web-design/bottom-up-visual-perception/

Brookes, J., Warburton, M., Alghadier, M., Mon-Williams, M., & Mushtaq, F. (2019). Studying human behavior with virtual reality: The Unity Experiment Framework. *Behavior Research Methods*, *52*, 455–463. doi.org/10.3758/s13428-019-01242-0

Brownlow, S., Dixon, A. R., Egbert, C. A., & Radcliffe, R. D. (1997). Perception of movement and dancer characteristics from point-light displays of dance. *The Psychological Record*, *47*, 411–421. https://link.springer.com/article/10.1007/BF03395235

Brownson, R. C., Eyler, A. A., Harris, J. K., Moore, J. B., & Tabak, R. G. (2018). Getting the Word Out: New approaches for disseminating public health science. *Journal of Public Health Management and Practice*, 24(2), 102-111. doi.org/10.1097/PHH.0000000000000673

Brunswik, E. (1952). The conceptual framework of psychology. *International Encyclopedia of Unified Science*,1(10). Chicago: University of Chicago Press. doi.org/10.1037/h0053448

Buckley, J. P. (2012). Detection of deception researchers needs to collaborate with experienced practitioners. *Journal of Applied Research in Memory and Cognition*, *1*(2), 126–127. doi.org/10.1016/j.jarmac.2012.04.002

Burgoon, J. K. (2015). Expectancy violations theory. *The International Encyclopedia of Interpersonal Communication*, 1–9. American Cancer Society. doi.org/https://doi.org/10.1002/9781118540190.wbeic102

Burgoon, J. K., Adkins, M., Kruse, J., Jensen, M. L., Meservy, T. O., Twitchell, D. P., Deokar, A., Nunamaker, J. F., Lu, S., Tsechpenakis, G., Metaxas, D. N., & Younger, R. E. (2005). An approach for intent identification by building on deception detection. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 1-10. doi.org/10.1109/hicss.2005.78

Busscher, B., Spinhoven, P., van Gerwen, L. J., & de Geus, E. J. C. (2013). Anxiety sensitivity moderates the relationship of changes in physiological arousal with flight anxiety during in vivo exposure therapy. *Behaviour Research and Therapy*, *51*(2), 98–105. doi.org/10.1016/j.brat.2012.10.009

Cambridge Dictionary (2020). Deception. *Cambridge English Dictionary*. Cambridge Univesity Press. https://dictionary.cambridge.org/dictionary/english/computer-aided

Cañal-Bruland, R., van der Kamp, J., & van Kesteren, J. (2010). An examination of motor and perceptual contributions to the recognition of deception from others' actions. *Human Movement Science*, *29*(1), 94–102. doi.org/10.1016/j.humov.2009.10.001

Canessa, A., Casu, P., Solari, F., & Chessa, M. (2019). Comparing real walking in immersive virtual reality and in physical world using gait analysis. *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 121–128. doi.org/10.5220/0007380901210128

Carson, T. (2001). Deception and withholding information in sales. *Business Ethics Quarterley*, *11*(2), 275–306.

Cerda, L., Fauvarque, A., Graziani, P., & Del-Monte, J. (2021). Contextual priming to increase the sense of presence in virtual reality: exploratory study. *Virtual Reality*, *2003*. doi.org/10.1007/s10055-021-00515-4

Chan, J. C. P., Leung, H., Tang, J. K. T., & Komura, T. (2011). A virtual reality dance training system using motion capture technology. *IEEE Transactions on Learning Technologies*, *4*(2), 187–195. doi.org/10.1109/TLT.2010.27

Chan, Z. Y. S., Macphail, A. J. C., Au, I. P. H., Zhang, J. H., Lam, B. M. F., Ferber, R., & Cheung, R. T. H. (2019). Walking with head-mounted virtual and augmented reality devices: Effects on position control and gait biomechanics. *Plos ONE*, *14*(12), 1–14. doi.org/10.1371/journal.pone.0225972

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*(4), 293–332.

Choudhury, T., & Fenwick, H. (2011). The impact of counter-terrorism measures on Muslim communities. *Equality and Human Rights Commission*. doi.org/10.1080/13600869.2011.617491

Cleary, H. M. D., & Warner, T. C. (2016). Police training in interviewing and interrogation methods: A Comparison of techniques used with adult and juvenile suspects. *Law and Human Behavior*, *40*(3), 270–284. doi.org/10.1037/lhb0000175

Cohen, C. J., Morelli, F., & Scott, K. A. (2008). A surveillance system for the recognition of intent within individuals and crowds. *2008 IEEE International Conference on Technologies for Homeland Security*. doi.org/10.1109/THS.2008.4534514

Cohen, D., Beattie, G., & Shovelton, H. (2010). Nonverbal indicators of deception: How iconic gestures reveal thoughts that cannot be suppressed. *Semiotica*, *2010*(182), 133–174. doi.org/10.1515/semi.2010.055

Committee on Science, Space. And Technology (2011). Behavioral Science and Security: Evaluating TSA's SPOT Program. *govinfo*. https://www.govinfo.gov/content/pkg/CHRG-112hhrg65053/pdf/CHRG-112hhrg65053.pdf

Davis, P., Perry, W., Brown, R., Yeung, D., Roshan, P., & Voorhies, P. (2013). Using behavioral indicators to detect potential violent acts. *RAND Corporation*. doi.org/10.7249/rb9724

Davis, R. (1961). Physiological responses as a means of evaluating information. *The Manipulation of Human Behavior*,142–168. Wiley. https://doc.uments.com/g-physiological-responses-as-a-means-of-evaluating-information.pdf

De Groot, J. H. B., Beetsma, D. J. V., van Aerts, T. J. A., le Berre, E., Gallagher, D., Shaw, E., Aarts, H., & Smeets, M. A. M. (2020). From sterile labs to rich VR: Immersive multisensory context critical for odors to induce motivated cleaning behavior. *Behavior Research Methods*, *52*(4), 1657–1670. doi.org/10.3758/s13428-019-01341-y

De Leva, P. (1996). Adjustments to Zatsiorsky-Seluyanov’s segment inertia parameters. *Journal of Biomechanics*, *29*(9), 1223–1230. doi.org/10.1002/ima.22019

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., Penke, L., Rafnsson, S. B., & Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin*, *92*(1), 135–152. doi.org/10.1093/bmb/ldp033

Decatoire, A., Beauprez, S. A., Pylouster, J., Lacouture, P., Blandin, Y., & Bidet-Ildei, C. (2019). Plavimop: How to standardize and simplify the use of point-light displays. *Behavior Research Methods*, *51*(6), 2573–2596. doi.org/10.3758/s13428-018-1112-x

Demott, R. (2010). *Xsens & the third floor help previsualize Iron Man 2*. *Animation World Network*. https://www.awn.com/news/xsens-third-floor-help-previsualize-iron-man-2

Denault, V. (2020). Misconceptions about nonverbal cues to deception: a covert threat to the justice system? *Frontiers in Psychology*, *11*, 1–4. doi.org/10.3389/fpsyg.2020.573460

Denault, V., & Jupe, L. M. (2018). Aviation security and the TSA’s behavior detection: Why effective academic and practitioner dialogue is vital. In *Frontiers in Psychology*, 9, 1-2.

doi.org/10.3389/fpsyg.2018.00240

Denault, V., Plusquellec, P., Jupe, L. M., St-Yves, M., Dunbar, N. E., Hartwig, M., Sporer, S. L., Rioux-Turcotte, J., Jarry, J., Walsh, D., Otgaar, H., Viziteu, A., Talwar, V., Keatley, D. A., Blandón-Gitlin, I., Townson, C., Deslauriers-Varin, N., Lilienfeld, S. O., Patterson, M. L., Koppen, P. J. V. (2020). The analysis of nonverbal communication: the dangers of pseudoscience in security and justice contexts. *Anuario de Psicologia Juridica*, *30*(1), 1–12. doi.org/10.5093/apj2019a9

Depaulo, B. M. (1992). Nonverbal behavior and self-presentation.(Comportamiento no verbal y auto-presentación). *Psychological Bulletin*, *111*(2), 203–243. https://doi.org/10.1037/0033-2909.111.2.203

Depaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*,*70*(5), 979-995.

Depaulo, B. M., Kirkendol, S. E., Tang, J., & O'Brien, T. P. (1988). The motivat!onal impairment effect in the communication of deception: Replications and extensions. *Journal of Nonverbal Behavior*, 12(3), 177-202. https://link.springer.com/content/pdf/10.1007%2FBF00987487.pdf

Depaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., Cooper, H., Lindsay, J. J., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*(1), 74–118. doi.org/10.1037/0033-2909.129.1.74

Derrick, D. C., Moffitt, K., & Nunamaker, J. F. (2011). Eye gaze behavior as a guilty knowledge test: Initial exploration for use in automated, kiosk-based screening. *Hawaii International Conference on System Sciences*, 4–7.

Dijksterhuis, A., & Nordgren, L. (2006). A theory of unconscious thought. *Perspectives on*

Psychological Science, *1*(2), 95–109. doi.org/10.1111/j.1745-6916.2006.00007.x

Dittrich, W. H., Trosciankoh, T., Lea, S. E. G., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, *25*(6).727-738 https://journals.sagepub.com/doi/pdf/10.1068/p250727

Dowler, K., Fleming, T., & Muzzatti, S. L. (2006). Constructing crime: Media, crime, and popular culture. *Canadian Journal of Criminology and Criminal Justice*, *48*(6), 837–850. doi.org/10.3138/cjccj.48.6.837

Dretske, F. (2006). Perception without awareness., *Perceptual experience*, 147–180. Oxford University Press. doi.org/10.1016/0001-6918(67)90013-3

Drillis, R., Contini, R., & Bluestein, M. (1964). Body segment parameters. A survey of measurement techniques. *Artificial Limbs*, *8*, 44–66. doi.org/10.1049/ecej:19890011

Durgin, F. H., Doyle, E., & Egan, L. (2008). Upper-left gaze bias reveals competing search strategies in a reverse Stroop task. *Acta Psychologica*, *127*(2), 428–448. doi.org/10.1016/j.actpsy.2007.08.007

Eapen, N., Baron, S., Street, C., & Richardson, D. (2010). The bodily movements of liars. *Cognitive Science Society*, *January 2010*, 1–7.

Ekman, P. (1988). Lying and non-verbal behaviour: Theoretical issues and new findings. *Journal of Nonverbal Behavior*, *12*, 163–175.

Ekman, P. (2011). *Behavioral detection officers & the SPOT programme*. Ekman International. https://www.ekmaninternational.com/behavioral-detection-officers-the-spot-programme/

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, *32*,88–106). doi.org/10.1080/00332747.1969.11023575

Ekman, P., Friesen, W. V., & Scherer, K. R. (1976). Body movement and voice pitch in deceptive interaction. *Semiotica*, *16*(1), 23–28. doi.org/10.1515/semi.1976.16.1.23

Erdfelder, E., Faul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. doi.org/10.3758/BRM.41.4.1149

Eubanks, D. L., Zhang, K., & Frumkin, L. (2015). Non-verbal cues to deception and their relationship to terrorism. *Hostile Intent and Counter-Terrorism: Human Factors Theory and Application*, 75–92). Ashgate Publishing. doi.org/10.1201/9781315587080-8

Ferrari, A., Cutti, A. G., Garofalo, P., Raggi, M., Heijboer, M., Cappello, A., & Davalli, A. (2010). First in vivo assessment of "outwalk": A novel protocol for clinical gait analysis based on inertial and magnetic sensors. *Medical and Biological Engineering and Computing*, *48*(1), 1–15. doi.org/10.1007/s11517-009-0544-y

Fink, P. W., Foo, P. S., & Warren, W. H. (2007). Obstacle Avoidance During Walking in Real and Virtual Environments. *ACM Transactions on Applied Perception*, *4*(1), 2. doi.org/10.1145/1227134.1227136

Fisher, R., Perkins, S., Walker, A., & Wolfart, E. (2003). *Spatial filters - median filter*. University of Edinburgh School of Informatics. https://homepages.inf.ed.ac.uk/rbf/HIPR2/median.htm

Ford, E. B. (2006). Lie detection: Historical, neuropsychiatric and legal dimensions. *International Journal of Law and Psychiatry*, *29*(3), 159–177. doi.org/10.1016/j.ijlp.2005.07.001

Frank, M. G., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stake lies. *Journal of Personality and Social Psychology*, 97 (6).

Frank, M. G., Menasco, M. A., O'Sullivan, M., & Voeller, J. G. (2008). Human behavior and deception detection. *Handbook of Science and Technology for Homeland Security*, 5. https://www.cl.cam.ac.uk/~rja14/shb10/frank2.pdf

Gillen, D., & Morrison, W. G. (2015). Aviation security: Costing, pricing, finance and performance. *Journal of Air Transport Management*, *48*, 1–12. doi.org/10.1016/j.jairtraman.2014.12.005

Gilovich, T., Savitsky, K., & Medvec, V. H. (1998). The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology*, *75*(2), 332–346. doi.org/10.1037/0022-3514.75.2.332

Gkritza, K., Niemeier, D., & Mannering, F. (2006). Airport security screening and changing passenger satisfaction: An exploratory assessment. *Journal of Air Transport Management*, *12*, 213–219. doi.org/10.1016/j.jairtraman.2006.03.001

Gorbunova, E. S., & Falikman, M. V. (2019). Visual search for letters in the right versus left visual hemifields. *Advances in Cognitive Psychology*, *15*(2), 75–88. doi.org/10.5709/acp-0258-5

Gorski, T. N. (1998). Defining deception. *Texas medicine* ,94(10), 9-10.

Grafton, S. T. (2009). Embodied cognition and the simulation of action to understand others. *Annals of the New York Academy of Sciences*, *1156*, 97–117. doi.org/10.1111/j.1749-6632.2009.04425.x

Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: Verbal and non-verbal cues to deception. *Applied Cognitive Psychology*, *16*(3), 243–257. doi.org/10.1002/acp.784

Granhag, P. A., Vrij, A., & Verschuere, B. (2015). Detecting deception: Current challenges and cognitive approaches. *Detecting deception: Current challenges and cognitive approaches.* Wiley-Blackwell.

Greene, J. O., Dan O'Hair, H., Cody, M. J., & Yen, C. (1985). Planning and control of behavior during deception. *Human Communication Research*, *11*(3), 335–364. doi.org/10.1111/j.1468-2958.1985.tb00051.x

Gross, S. R. (2001). Detection of Deception : The case of handwriting expertise. *Virginia Law Review*, *87*(8), 1847–1855.

Grubin, D., & Madsen, L. (2007). A historical review. *The Journal of Forensic Psychiatry & Psychology*, *16*(2), 357–369. https://doi.org/10.1080/14789940412331337353

Guo, K., Meints, K., Hall, C., Hall, S., & Mills, D. (2009). Left gaze bias in humans, rhesus monkeys and domestic dogs. *Animal Cognition*, *12*(3), 409–418. doi.org/10.1007/s10071-008-0199-3

Gustafsson, P. U., Lindholm, T., & Jönsson, F. U. (2019). Predicting accuracy in eyewitness testimonies with memory retrieval effort and confidence. *Frontiers in Psychology*, *10,* 1-10. doi.org/10.3389/fpsyg.2019.00703

Hainen, A. M., Remias, S. M., Bullock, D. M., & Mannering, F. L. (2013). A hazard-based analysis of airport security transit times. *Journal of Air Transport Management*, *32*, 32–38. doi.org/10.1016/j.jairtraman.2013.06.002

Harcourt, B. E. (2011). Against prediction: Sentencing, policing, and punishing in an actuarial age. *Chicago Public Law and Legal Theory: University of Chicago Press*, *1-43*. doi.org/10.2139/ssrn.756945

Harris, D. H. (2002). How to really improve airport security. *Ergonomics in Design,1-6*.

Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin*, *137*(4), 643–659. doi.org/10.1037/a0023589

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law and Human Behavior*, *30*, 603–619. doi.org/10.1007/s10979-006-9053-9

Hartwig, M., Granhag, P. A., Strömwall, L. A., Wolf, A. G., Vrij, A., & Roos af Hjelmsäter, E. (2011). Detecting deception in suspects: Verbal cues as a function of interview strategy. *Psychology, Crime and Law*, *17*(7), 643–656. doi.org/10.1080/10683160903446982

Hashad, D. (2004). Stolen freedoms: Arabs, muslims, and south asians in the wake of post 9/11 backlash. *Denver University Law Review*, *81*(4), 735–748. http://eprints.uanl.mx/5481/1/1020149995.PDF

Hayes, R. M., & Levett, L. M. (2013). Community members' perceptions of the CSI effect. *American Journal of Criminal Justice*, *38*(2), 216–235. https://doi.org/10.1007/s12103-012-9166-2

Herbert, L. (2007). Othello error: Facial profiling, privacy, and the suppression of dissent. *Ohio State Journal of Criminal Law*, *5*, 79–129.

Hicks, C., & Ulvestad, N. (2011). Deception detection accuracy using verbal or nonverbal cues. *The Journal of Undergraduate Research, 9, 63-68*. https://openprairie.sdstate.edu/jur/vol9/iss1/9

Holder, M. K. (2001). *Why are more people right-handed? - Scientific American*. https://www.scientificamerican.com/article/why-are-more-people-right/

Home Office (2020). Mandatory polygraph tests factsheet. *GOV.UK.* https://www.gov.uk/government/publications/domestic-abuse-bill-2020-factsheets/mandatory-polygraph-tests-factsheet

Homeland Security Digital Library. (2002). Aviation security with a focus on passenger profiling. *Homeland Security Digital Library, 1-59*.

Horvath, F. S., & Reid, J. E. (1971). The reliability of polygraph examiner diagnosis of truth and deception. *The Journal of Criminal Law, Criminology, and Police Science*, *62*(2), 276. doi.org/10.2307/1141892

Hurrell, K. (2013). *An experimental analysis of examinations and detentions under schedule 7 of the terrorism act 2000*.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., & Mazziotta, J. C. (2005). Grasping the intentions of others with one's own mirror neuron system. *Plos Biology*, *3*(3), 529–535. doi.org/10.1371/journal.pbio.0030079

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). Criminal Interrogation and Confessions. Aspen Publishers.

CPNI (2020). Behavioural detection. *Centre for the Protection of National Infrastructure.* https://www.cpni.gov.uk/behavioural-detection-0

Innocence Project. (2019). Central Park Five Tragedy Reframed in Netflix Series "When They See Us.". *Innocence Project*. https://innocenceproject.org/central-park-five-tragedy-reframed-in-netflix-series-when-they-see-us/

Janeh, O., Langbehn, E., Steinicke, F., Bruder, G., Gulberti, A., & Poetter-Nerger, M. (2017). Walking in virtual reality: Effects of manipulated visual self-motion on walking biomechanics. *ACM Transactions on Applied Perception*, *14*(2), 1–15. doi.org/10.1145/3022731

Jupe, L. M., & Denault, V. (2019). Science or pseudoscience? A distinction that matters for police officers, lawyers and judges. *Psychiatry, Psychology and Law*, *26*(5), 753–765. doi.org/10.1080/13218719.2019.1618755

Jupe, L. M., & Keatley, D. A. (2019). Airport artificial intelligence can detect deception: or am I lying? *Security Journal*, 1–26. doi.org/10.1057/s41284-019-00204-7

Kahneman, D. (1991). Judgment and decision making: A personal view. *Psychological Science, 2*(3), 142–145. doi.org/10.1111/j.1467-9280.1991.tb00121.x

Kelly, C. (2009). The 'Israelification' of airports : High security , little bother. *The Star*. https://www.thestar.com/news/world/2009/12/30/the_israelification_of_airports_high_security_little_bother.html

Kippenberg, H. G., & Seidensticker, T. (2006). The 9/11 Handbook- Annotated translation and interpretation of the attackers' spiritual manual. Equinox.

Kozinski, W. (2018). The Reid interrogation technique and false confessions : A time for change. *Seattle Journal for Social Justice*, *16*(2), 301–345.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, *21*(6), 575–580. https://link.springer.com/content/pdf/10.3758%2FBF03198740.pdf

Kuhlmeier, V. A., Troje, N. F., & Lee, V. (2010). Young infants detect the direction of biological motion in point-light displays. *Infancy*, *15*(1), 83–93. doi.org/10.1111/j.1532-7078.2009.00003.x

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). Lmertest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13). doi.org/10.18637/jss.v082.i13

Laviola Jr., J. J., Kruijff, E., Bowman, D. A., McMahan, R. P., & Poupyrev, I. (2017). 3D interaction techniques. *3D User Interfaces: Theory and Practice*. Addison-Wesley Professional. https://learning.oreilly.com/library/view/3d-user interfaces/9780134034478/pref04.html#copy

Lawson, G., Stedmon, A. W., Zhang, C., Eubanks, D. L., & Frumkin, L. (2013). Deception and self-awareness. *Frontiers in Psychology*, *4*, 1–16. doi.org/10.1007/978-3-642-21741-8_44

Lee, A. J., & Jacobson, S. H. (2011). The impact of aviation checkpoint queues on optimizing security screening effectiveness. *Reliability Engineering and System Safety*. doi.org/10.1016/j.ress.2011.03.011

Leo, R. A. (2009). False confessions: Causes, consequences, and implications. *Journal of the American Academy of Psychiatry and the Law*, *37*(3), 332–343.

Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, *33*(4), 378–392. doi.org/10.1177/0261927X14535916

Levine, T. R. (2018). Ecological validity and deception detection research design. *Communication Methods and Measures*, *12*(1), 45–54. doi.org/10.1080/19312458.2017.1411471

Levine, T. R., Asada, K. J. K., & Park, H. S. (2006). The lying chicken and the gaze avoidant egg: Eye contact, deception, and causal order. *Southern Communication Journal*, *71*(4), 401–411. doi.org/10.1080/10417940601000576

Levine, T. R., Serota, K. B., & Shulman, H. C. (2010). The impact of lie to me on viewers' actual ability to detect deception. *Communication Research*, *37*(6), 847–856. doi.org/10.1177/0093650210362686

Lewis, J., & Marsden, S. (2020). Public Experiences of the UK Counter-Terrorism System. *Centre for Research and Evidence on Security Threats.* https://crestresearch.ac.uk/resources/public-experiences-of-the-uk-counter-terrorism-system/

Lilienfeld, S. O. (2017). Psychology's replication crisis and the grant culture: Righting the Ship. *Perspectives on Psychological Science*, *12*(4), 660–664. doi.org/10.1177/1745691616687745

Lilienfeld, S. O., & Landfield, K. (2008). Science and pseudoscience in law enforcement: A user-friendly primer. *Criminal Justice and Behavior*, *35*(10), 1215–1230. doi.org/10.1177/0093854808321526

Limbasiya, H. (2018). Sense simulation in virtual reality to increase: Immersion, presence and interactions. *University of Dublin*. https://www.scss.tcd.ie/publications/theses/diss/2018/TCD-SCSS-DISSERTATION-2018-066.pdf

Lindholm, T., Jönsson, F. U., & Liuzza, M. T. (2018). Retrieval effort cues predict eyewitness accuracy. *Journal of Experimental Psychology: Applied*, *24*(4), 534–542. doi.org/10.1037/xap0000175

Loftus, E. F., Loftus, G. R., & Messo, J. (1987). Some facts about "weapon focus." *Law and Human Behaviour*, *11*(1), 55–62.

Lu, S., Tsechpenakis, G., Metaxas, D. N., Jensen, M. L., & Kruse, J. (2005). Blob analysis of the head and hands: A method for deception detection. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 20. doi.org/10.1109/hicss.2005.122

Luke, T. J. (2019). Lessons from pinocchio: Cues to deception may be highly exaggerated. *Perspectives on Psychological Science*, *14*(4), 646–671. doi.org/10.1177/1745691619838258

Mahon, J. E. (2016). The definition of lying and deception. *The Stanford Encyclopedia of Philosophy*,1–14. https://plato.stanford.edu/entries/lying-definition/

Mandelbaum, J. (2014). Exploring The Illusion Of Transparency When Lying And Truth-Telling: The Impact Of Age, Self-Consciousness, And Framing. *City University of New York,* 1-167. https://academicworks.cuny.edu/cgi/viewcontent.cgi?article=1504&context=gc_etds

Mapala, T., Warmelink, L., & Linkenauger, S. A. (2017). Jumping the gun: Faster response latencies to deceptive questions in a realistic scenario. *Psychonomic Bulletin & Review*, *24*(4), 1350–1358. doi.org/10.3758/s13423-016-1218-z

Marshall, D., & Thomas, T. (2015). Polygraphs and sex offenders: The truth is out there. *Probation Journal*, *62*(2), 128–139. doi.org/10.1177/0264550515571395

Martin, F. (2018). Airport security: Trying to prevent the worst from occurring while protecting from its deadly consequences. *Journal of Airport Management*, *12*(2), 144–156.

Masip, J., & Herrero, C. (2017). Examen del sesgo de respuesta de los policías al juzgar la veracidad. *Psicothema*, *29*(4), 490–495. doi.org/10.7334/psicothema2016.357

Matsumoto, D., Hwang, H., Fullenkamp, A., & Laurent, C. (2015). Human deception detection from whole body motion analysis. *Air Force Research Laboratory,* 1-55. https://apps.dtic.mil/dtic/tr/fulltext/u2/a626755.pdf

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis?: What does “failure to replicate” really mean? *American Psychologist*, *70*(6), 487–498. doi.org/10.1037/a0039400

Meijer, E. H., & Verschuere, B. (2010). The polygraph and the detection of deception. *Journal of Forensic Psychology Practice*, *10*(4), 325–338. doi.org/10.1080/15228932.2010.481237

Meyer, D. L. (2010). The SPOT program : Hello racial profiling, goodbye fourth amendment ? *University of Maryland Law Journal of Race, Religion, Gender and Class*, *10*(2).

Michael, N., Dilsizian, M., Metaxas, D. N., & Burgoon, J. K. (2010). Motion profiles for deception detection using visual cues. *Lecture Notes in Computer Science*, 462–475. doi.org/10.1007/978-3-642-15567-3_34

Modić, A., Vidović, A., & Lovrić, Ž. (2018). Airport landside security: From recommended practice to standard. *18th International Conference on Transport Science*.

Moi, W. Y., & Shanks, D. R. (2015). Can lies be detected unconsciously? *Frontiers in Psychology*, *6*, 1–10. doi.org/10.3389/fpsyg.2015.01221

Mullin, D. S., King, G. W., Saripalle, S. K., Derakhshani, R. R., Lovelace, C. T., & Burgoon, J. K. (2014). Deception effects on standing center of pressure. *Human Movement Science*, *38*, 106–115. doi.org/10.1016/j.humov.2014.08.009

Murai, J., Nose, I., & Takiguchi, Y. (2018). The more attractive, the less deceptive? Effects of female facial attractiveness on perceived deceptiveness. *Psychology*, *9*(4), 529–539. doi.org/10.4236/psych.2018.94032

MVN (2015). MVN User Manual. https://www.xsens.com/hubfs/Downloads/usermanual/MVN_User_Manual.pdf

Nilsson, N. C., Serafin, S., & Nordahl, R. (2016). The perceived naturalness of virtual walking speeds during WIP locomotion: Summary and meta-analyses. *Psychology Journal*, *14*(1), 7–39.

Nilsson, N. C., Serafin, S., Steinicke, F., & Nordahl, R. (2018). Natural walking in virtual reality: A review. *Computers in Entertainment*, *16*(2). doi.org/10.1145/3180658

Norman, D. G., Wade, K. A., Williams, M. A., & Watson, D. G. (2020). Caught virtually lying-Crime scenes in virtual reality help to expose suspects' concealed recognition. *Journal of Applied Research in Memory and Cognition*, *9*, 118–127.

Nortje, A., & Tredoux, C. (2019). How good are we at detecting deception? A review of current techniques and theories. *South African Journal of Psychology*, *49*(4), 491–504. doi.org/10.1177/0081246318822953

O'Sullivan, M., Frank, M. G., Hurley, C. M., & Tiwana, J. (2009). Police lie detection accuracy: The effect of lie scenario. *Law and Human Behavior*, *33*, 530–538. doi.org/10.1007/s10979-008-9166-4

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. doi.org/10.1016/0028-3932(71)90067-4

Olivier, A. H., Bruneau, J., Kulpa, R., & Pettre, J. (2018). Walking with virtual people: Evaluation of locomotion interfaces in dynamic environments. *IEEE Transactions on Visualization and Computer Graphics*, *24*(7), 2251–2263. doi.org/10.1109/TVCG.2017.2714665

Omar, A., & Miah, M. (2014). Risk based airport security approach department of computer information systems. *International Journal of Computer Science: Theory and Applications*.

Ormerod, T. C., & Dando, C. J. (2014). Finding a needle in a haystack: toward a psychologically informed method for aviation security screening. *Journal of Experimental Psychology: General*.

Pak, J., & Zhou, L. (2013). Eye gazing behaviors in online deception. *19th Americas Conference on Information Systems*. https://core.ac.uk/download/pdf/301359386.pdf

Park, S. H., Ryu, D., Waiora, T. H., Abernethy, B., & Mann, D. L. (2019). Falling for a fake: The role of kinematic and non-kinematic information in deception detection liis uiga and rich masters. *Perception*, *48*(4), 330–337. doi.org/10.1177/0301006619837874

Parker, E. (2007). Implementation of the uk terrorism act 2006 the relationship between counterterrorism law, free speech, and the muslim community in the united kingdom versus the united states. *Emory International Law Review*, *21*(1), 711–758. doi.org/10.3868/s050-004-015-0003-8

Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in Human Neuroscience*, *9*, 1–19. doi.org/10.3389/fnhum.2015.00660

Pinar Saygin, A., Wilson, S. M., Hagler Jr, D. J., Bates, E., & Sereno, M. I. (2004). Point-light biological motion perception activates human premotor cortex. *The Journal of Neuroscience*, *24*, 6181–6188. doi.org/10.1523/JNEUROSCI.0504-04.2004

Pinizzotto, A. J., Davis, E. F., & Miller, C. E. (2004). Intuitive policing: Emotional/rational decision making in law enforcement. *FBI Law Enforcement Bulletin*, *73*, 1–6.

Pollard, K. A., Oiknine, A. H., Files, B. T., Sinatra, A. M., Patton, D., Ericson, M., Thomas, J., & Khooshabeh, P. (2020). Level of immersion affects spatial learning in virtual environments: results of a three-condition within-subjects study with long intersession intervals. *Virtual Reality*,*24*,783-796. https://link.springer.com/article/10.1007/s10055-019-00411-y

Poppe, R., Van Der Zee, S., Heylen, D. K. J., & Taylor, P. J. (2014). AMAB: Automated measurement and analysis of body motion. *Behavior Research Methods*, *46*(3), 625–633. doi.org/10.3758/s13428-013-0398-y

Poppe, R., Van Der Zee, S., Taylor, P. J., Anderson, R., & Veltkamp, R. C. (2014). Mining Bodily Cues to Deception, *Conference Proceedings of the Rapid Screening Technologies, Deception Detection and Credibility Assessment Symposium*. https://www.cl.cam.ac.uk/~rja14/Papers/mining-bodily-cues.pdf

Press, W. (2010). To catch a terrorist: Can ethnic profiling work? *Significance*, *7*(4), 164–167. doi.org/10.1111/j.1740-9713.2010.00452.x

Randhavane, T., Bhattacharya, U., Kapsaskis, K., Gray, K., Bera, A., & Manocha, D. (2019). The liar's walk detecting deception with gait and gesture. *Computer Science*. https://arxiv.org/pdf/1912.06874.pdf

Razzaque, S., Swapp, D., Slater, M., Whitton, M. C., & Steed, A. (2002). Redirected walking in place. *Proceedings of the Workshop on Virtual Environments, 123-130*.

Reinhard, M. A., Greifeneder, R., & Scharmach, M. (2013). Attitudes and social cognition: Unconscious processes improve lie detection. *Journal of Personality and Social Psychology*, *105*(5), 721–739. doi.org/10.1037/a0034352

Reynolds, E., & Rendle-Short, J. (2011). Cues to deception in context: Response latency/gaps in denials and blame shifting. *British Journal of Social Psychology*, *50*(3), 431–449. doi.org/10.1348/014466610X520104

Rstudio (2015). Integrated development for r. rstudio. *Rstudio, Inc., Boston*. http://www.rstudio.com/.

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, *112*(4), 585–615. doi.org/10.1037/0096-3445.112.4.585

Sakano, R., Obeng, K., & Fuller, K. (2016). Airport security and screening satisfaction: A case study of U.S. *Journal of Air Transport Management*, *55*, 129–138. doi.org/10.1016/j.jairtraman.2016.05.007

Sartori, G., Orru, G., & Monaro, M. (2016). Detecting deception through kinematic analysis of hand movement. *International Journal of Psychophysiology, 108*, 16. doi.org/10.1016/j.ijpsycho.2016.07.056

Schuetzler, R. M. (2012). Countermeasures and eye tracking deception detection. *Information Systems and Quantitative Analysis Faculty Proceedings & Presentations*, 1–6.

Serchi, V., Peruzzi, A., Cereatti, A., & Della Croce, U. (2016). Use of a remote eye-tracker for the analysis of gaze during treadmill walking and visual stimuli exposition. *Biomed Research International*, *2016*. doi.org/10.1155/2016/2696723

Serota, K. B., & Levine, T. R. (2015). A few prolific liars. *Journal of Language and Social Psychology*, *34*(2), 138–157. doi.org/10.1177/0261927X14528804

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in america: Three studies of self-reported lies. *Human Communication Research*, *36*(1), 2–25. doi.org/10.1111/j.1468-2958.2009.01366.x

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, *69*, 487–510. doi.org/10.1146/annurev-psych-122216-011845

Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *3*(2), 130–144. doi.org/10.1162/pres.1994.3.2.130

Soczawa-Stronczyk, A. A., & Bocian, M. (2020). Gait coordination in overground walking with a virtual reality avatar: Gait coordination in virtual walking. *Royal Society Open Science*, *7*(7). doi.org/10.1098/rsos.200622

Lo, S., & Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology,6,* 1-16. doi.org/10.3389/fpsyg.2015.01171

Sporer, S. L., Masip, J., & Cramer, M. (2014). Guidance to detect deception with the aberdeen report judgment scales: Are Verbal content cues useful to detect false accusations? *The American Journal of Psychology*, *127*(1), 43–61. doi.org/10.5406/amerjpsyc.127.1.0043

Sporer, S. L., & Schwandt, B. (2006). Paraverbal Indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, *20*, 421–446. doi.org/10.1002/acp.1190

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law*, *13*(1), 1–34. doi.org/10.1037/1076-8971.13.1.1

Stein, J. P., & Ohler, P. (2017). Venturing into the uncanny valley of mind-The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition*, *160*, 43–50. doi.org/10.1016/j.cognition.2016.12.010

Stel, M., Schwarz, A., van Dijk, E., & van Knippenberg, A. (2020). The limits of conscious deception detection: When reliance on false deception cues contributes to inaccurate judgments. *Frontiers in Psychology*, *11*, 1–12. doi.org/10.3389/fpsyg.2020.01331

Stewart, M. G., & Mueller, J. (2014). Cost-benefit analysis of airport security: Are airports too safe? *Journal of Air Transport Management*, *35*, 19–28. doi.org/10.1016/j.jairtraman.2013.11.003

Stotz, T., Bearth, A., Ghelfi, S. M., & Siegrist, M. (2020). Evaluating the perceived efficacy of randomized security measures at airports. *Risk Analysis*, *40*(7), 1469–1480. doi.org/10.1111/risa.13474

Street, C. N. H., & Vadillo, M. A. (2016). Can the unconscious boost lie detection accuracy? *Current Directions in Psychological Science*, *25*(4), 246–250.

Ströfer, S., Noordzij, M. L., Ufkes, E. G., & Giebels, E. (2015). Deceptive intentions: Can cues to deception be measured before a lie is even stated? *Plos ONE*, *10*(5). doi.org/10.1371/journal.pone.0125237

Ströfer, S., Ufkes, E. G., Noordzij, M. L., & Giebels, E. (2016). Catching a deceiver in the act: Processes underlying deception in an interactive interview setting. *Applied Psychophysiology Biofeedback*, *41*(3), 349–362. doi.org/10.1007/s10484-016-9339-8

Strömwall, L. A., & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime and Law*, *9*(1), 19–36. doi.org/10.1080/10683160308138

Suchotzki, K., Verschuere, B., Crombez, G., & De Houwer, J. (2013). Reaction time measures in deception research: Comparing the effects of irrelevant and relevant stimulus-response compatibility. *Acta Psychologica*. doi.org/10.1016/j.actpsy.2013.06.014

Suma, E. A., Azmandian, M., Grechkin, T., Phan, T., & Bolas, M. (2015). Making small spaces feel large: Infinite walking in virtual reality. *SIGGRAPH 2015 Emerging Technologies*, 1. doi.org/10.1145/2782782.2792496

The Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology*, *37*(1), 60–74. doi.org/10.1177/0022022105282295

Thompson, J. (2012). Thin slices & first impressions. *Psychology Today*. https://www.psychologytoday.com/gb/blog/beyond-words/201203/thin-slices-first-impressions

Tobiipro. (2015). *How do Tobii eye trackers work? - Learn more with Tobii Pro*. *TobiiPro*. https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/how-do-tobii-eye-trackers-work/

Twyman, N. (2012). Automated human screening for detecting concealed knowledge. The

University of Arizona. https://scholarsmine.mst.edu/cgi/viewcontent.cgi?article=1177&context=bio_inftec_facwork

Twyman, N., Elkins, A. C., Burgoon, J. K., & Nunamaker, J. F. (2014). A rigidity detection system for automated credibility assessment. *Journal of Management Information Systems*, *31*(1), 173–202. doi.org/10.2753/MIS0742-1222310108

U.S. Government Accountability Office (2010). Efforts to validate tsa's passenger screening behavior detection program underway, but opportunities exist to strengthen validation and address operational challenges. *gao.gov*. https://www.gao.gov/products/gao-10-763

U.S. Government Accountability Office (2017). Aviation security: TSA does not have valid evidence supporting most of the revised behavioral indicators used in its behavior detection activities. *gao.gov*. https://www.gao.gov/products/GAO-17-608R

Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments*, *9*(5), 497–503. doi.org/10.1162/105474600566989

Van Der Zee, S., Poppe, R., Taylor, P. J., & Anderson, R. (2015). To freeze or not to freeze A motion-capture approach to detecting deceit. *Proceedings of the Hawaii International Conference on System Sciences*, 1–18.

Verrel, J., Lövdén, M., Schellenbach, M., Schaefer, S., & Lindenberger, U. (2009). Interacting effects of cognitive load and adult age on the regularity of whole-body motion during treadmill walking. *Psychology and Aging*, *24*(1), 75–81. doi.org/10.1037/a0014272

Vicianova, M., & Jozef, P. (2015). Literature reviews historical techniques of lie detection. *Europe's Journal of Psychology*, *11*(3), 522–534. doi.org/10.5964/ejop.v11i3.919

Vrij, A. (1995). Behavioral correlates of deception in a simulated police interview. *The Journal of Psychology*, *129*(1), 15–28.

Vrij, A. (2000). Beliefs about Nonverbal and Verbal Cues to Deception. *Detecting Lies and Deceit.* Wiley.

Vrij, A. (2008a). *Detecting lies and deceit pitfalls and opportunities*. Wiley.

Vrij, A. (2008b). Nonverbal dominance versus verbal accuracy in lie detection: A plea to change police practice. *Criminal Justice and Behavior*, *35*(10), 1323–1336. doi.org/10.1177/0093854808321530

Vrij, A., & Fisher, R. P. (2020). Unraveling the misconception about deception and nervous behavior. *Frontiers in Psychology*, *11,* 1-8. doi.org/10.3389/fpsyg.2020.01377

Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology*, *22*, 1-21. doi.org/10.1111/lcrp.12088

Vrij, A., Fisher, R. P., Mann, S., & Leal, S. (2008). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling*, *5*, 39–43. doi.org/10.1002/jip.82

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Lying about flying: The first experiment to detect false intent. *Psychology, Crime & Law*, *17*, 611–620. doi.org/10.1080/10683160903418213

Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest*, *11*(3), 89–121. doi.org/10.1177/1529100610390861

Vrij, A., Hartwig, M., & Granhag, P. A. (2019). Reading lies: Nonverbal communication and deception. *Annual Review of Psychology*, *70*, 295–317.

https://www.annualreviews.org/doi/abs/10.1146/annurev-psych-010418-103135

Vrij, A., Mann, S. A., Ronald, A. E., Ae, P. F., Leal, S., Rebecca, A. E., Ae, M., Bull, R., Vrij, A., Mann, S. A., Leal, Á. S., Milne, Á. R., Fisher, R. P., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behaviour*, *32*, 253–265. doi.org/10.1007/s10979-007-9103-y

Vrij, A., Meissner, C. A., & Kassin, S. M. (2015). Problems in expert deception detection and the risk of false confessions: No proof to the contrary in Levine et al. (2014). *Psychology, Crime & Law*, *21*(9), 901–909. doi.org/10.1080/1068316X.2015.1054389

Vrij, A., & Semin, G. R. (1996). Lie experts' beliefs about nonverbal indicators of deception. *Journal of Nonverbal Behavior*, *20*, 65–80.

Vrij, A., Semin, G. R., & Bull, R. (1996). Insight into behavior displayed during deception. *Human Communication Research*, *22*(4), 544–562. doi.org/10.1111/j.1468-2958.1996.tb00378.x

Vrij, A., & Winkel, F. W. (1991). Cultural patterns in Dutch and Surinam nonverbal behavior: An analysis of simulated police/citizen encounters. *Journal of Nonverbal Behavior*, *15*(3), 169–184. doi.org/10.1007/BF01672219

Walczyk, J. J., Igou, F. P., Dixon, A. P., & Tcholakian, T. (2013). Advancing lie detection by inducing cognitive load on liars: A review of relevant theories and techniques guided by lessons from polygraph-based approaches. *Frontiers in Psychology*, *4*, 1-13. doi.org/10.3389/fpsyg.2013.00014

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying to questions: Response time as a cue to deception. *Applied Cognitive Psychology*, *17*(7), 755–774. doi.org/10.1002/acp.914

Weinberger, S. (2010). Airport security: Intent to deceive? *Nature*, *465*, 412–415. doi.org/10.1038/465412a

Wertheimer, M., & Riezler, K. (1944). Gestalt theory. *Social Research*, *11*(1), 78–99.

Wigginton, M., Jensen, C. J., Graves, M., & Vinson, J. (2014). What is the role of behavioral analysis in a multilayered approach to aviation security? *Journal of Applied Security Research*, *9*(4), 393–417. doi.org/10.1080/19361610.2014.942828

Winter, J., & Currier, C. (2015). Exclusive: TSA's secret behavior checklist to spot terrorists. *The Intercept*. https://theintercept.com/2015/03/27/revealed-tsas-closely-held-behavior-checklist-spot-terrorists/

Wright, M. J., Bishop, D. T., Jackson, R. C., & Abernethy, B. (2013). Brain regions concerned with the identification of deceptive soccer moves by higher-skilled and lower-skilled players. *Frontiers in Human Neuroscience*, *7*, 1–15. doi.org/10.3389/fnhum.2013.00851

Wu, G., & Cavanagh, P. R. (1995). ISB recommendations for standardization in the reporting of kinematic data. *Journal of Biomechanics*, *28*(10), 1257–1261. doi.org/10.1016/0021-9290(95)00017-C

Wu, S., Mei, H., & Yan, J. (2019). Don't think carefully? Re-examining the effect of unconscious thought on deception detection. *Frontiers in Psychology*, *10*(APR). doi.org/10.3389/fpsyg.2019.00893

Yu, L., Zhang, Q., Hu, C., Huang, Q., Ye, M., & Li, D. (2015). Effects of different frequencies of rhythmic auditory cueing on the stride length, cadence, and gait speed in healthy young females. In *Journal of Physical Therapy Science*,27(2), 485-487. doi.org/10.1589/jpts.27.485

Zedner, L. (2013). The hostile border: Crimmigration, counter-terrorism, or crossing the line on rights? *New Criminal Law Review*, *22*(3), 328–345.

Zhang, K., Eubanks, D. L., Frumkin, L., Saikayasit, R., Stedmon, A. W., & Lawson, G. (2013). Telling the difference between deceiving and truth telling: An experiment in a public space. *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, 2013,* 1–8. doi.org/10.1109/FG.2013.6553800

Ziegler, M. G. (2004). Psychological stress and the autonomic nervous system. *Primer On The Autonomic Nervous System,*189–190. Elsevier.

Zloteanu, M. (2016). Emotions and Deception Detection. *University College London,* 1-352. https://discovery.ucl.ac.uk/id/eprint/1537296/1/Thesis%20-%20%20Mircea%20Zloteanu%20-%202016%20-%20PostViva.pdf

Zuckerman, M., Depaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology*, *14,*1-59. doi.org/10.1016/S0065-2601(08)60369-X

Appendix A- Avatar samples

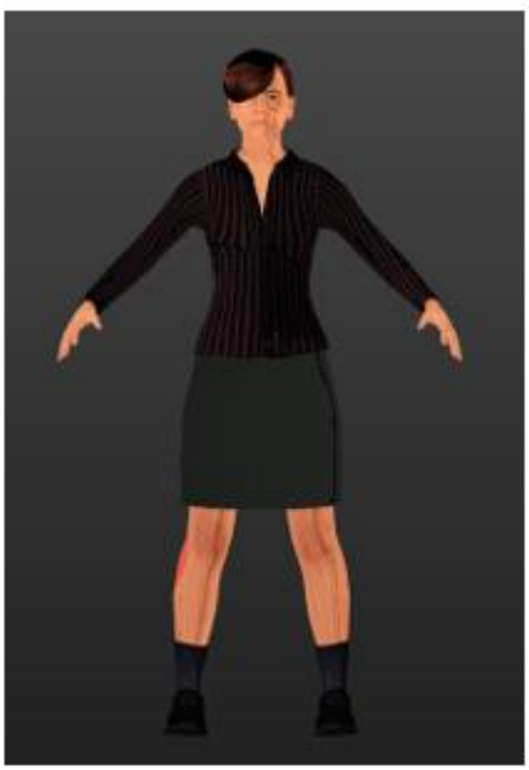
The images in this appendix are a small sample of the type of avatars that were used. Avatars are in the t-pose, which is not reflective of their posture when participants viewed them.












Appendix B- Immersion Questionnaire

Adapted from Slater, M., Usoh, M., & Steed, A. (1994).

1) Rate your sense of being at the security checkpoint, on the following scale from 1 to 7, where 7 represents your normal experience of being in a place.

   I had a sense of “being there” at the security checkpoint:

1	2	3	4	5	6	7
Not at all						Very much

2) To what extent were there times during the experience when the security checkpoint was the reality for you?

   There were times during the experience when the checkpoint was the reality for me...

1	2	3	4	5	6	7
At no time						All the time

3) When you think back about your experience, do you think of the security checkpoint more as images that you saw or more as somewhere that you visited?

   The security checkpoint seems to me to be more like...

1	2	3	4	5	6	7
Images that I saw						Somewhere that I visited

4) During the time of the experience, which was strongest, on the whole, your sense of being at the security checkpoint, or of being elsewhere?

I had a stronger sense of...

1	2	3	4	5	6	7
Being elsewhere						Being there

5) Consider your memory of being at the security checkpoint. How similar in terms of the structure of the memory is this to the structure of the memory of other places you have been today? By 'structure of the memory' consider things like the extent to which you have a visual memory of the checkpoint, whether that memory is in colour, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic in your imagination, and other such structural elements.

I think of the security checkpoint as a place in a way similar to other places that I've been today...

1	2	3	4	5	6	7
Not at all						Very much so

6) During the time of the experience, did you often think to yourself that you were actually at the security checkpoint?

During the experience I often thought that I was really at the security checkpoint...

1	2	3	4	5	6	7
Not very often						All the time

7) Please write down any further comments that you wish to make about your experience. In particular, what things helped to give you a sense of 'really being' at the security checkpoint, and what things acted to 'pull you out' of this?

Appendix C- Questions asked in Chapter Four

1. Did you pack your bag yourself?
2. Are you sure there are no prohibited items in your bag?
3. Do you have any illegal substances in your bag?
4. Do you have any electrical equipment in your bag?
5. Are there any unidentifiable objects in your bag?
6. Are you a student?
7. Are you travelling alone today?
8. Are all the items that you have within the travel guidelines?
9. Do you have any liquids over 100ml?
10. Do you have any explosive objects in your bag?
11. Are we currently in an airport?
12. Do you have any questionable or illegal items?
13. Are you meeting with anyone else before boarding?
14. Do you have any items which could be used to harm someone?
15. Do you have any prohibited items?
16. Do you have any tobacco products with you?