Investigating Conscious, Psychophysiological, and Behavioural Measures of Covert Surveillance Detection via Nonconventional Means

ROSS FRIDAY

A thesis submitted in partial fulfilment of the requirements of the University of Greenwich for the Degree of Doctor of Philosophy

May 2019

DECLARATION

"I certify that the work contained in this thesis, or any part of it, has not been accepted in substance for any previous degree awarded to me, and is not concurrently being submitted for any degree other than that of Doctor of Philosophy being studied at the University of Greenwich. I also declare that this work is the result of my own investigations, except where otherwise identified by references and that the contents are not the outcome of any form of research misconduct."

SIGNED: ____________________________

DATE: ____________________________

ACKNOWLEDGMENTS

Having taken the path of most resistance in return for the opportunity to contribute to a fascinating and important area of research, I would like to thank the organisations who helped make this possible. Without the grants awarded to me by the BIAL Foundation, the SPR and the PA, this now-completed thesis may still be a work in progress. This feels all the more meaningful to have taken place under the enigmatic guidance of Dr David Luke, with whom I discussed the seemingly remote possibility of starting an undergraduate degree during a chance meeting at a university open day almost a decade ago. What began as an instantly apparent shared sense of humour quickly grew into a joint pursuit of answers to science's difficult questions, and he has encouraged and inspired me at every stage to strive relentlessly for the next.

Along with fellow supervisors Dr Claire Monks and Dr Trevor Thompson, as well as hugely supportive colleagues such as Professor Chris Roe, Dr Richard Broughton, Dr James Demetre and Dr Josh Davis - David has helped me throughout an educational adventure which saw me go from being the first in my family to attend university, to achieving a doctorate. Ironically, I cannot help but feel my non-academic background played an essential role, and that being a carpenter's son is much of the reason I have completed a challenge that many do not. I have certainly needed the morals and values instilled in me throughout my amazing childhood when growing up with my hard-working parents and brother; it may have taken me years to fully appreciate the lessons they taught me, but I am certainly grateful now and can never thank them enough. I only wish that my much-missed Nan could have seen me graduate, as she would have been proven right having told anyone who would listen that I would finish what I started.

It is with mixed emotions that I conclude the enormous task that has loomed over me for so long, but it is time for another chapter. I arrived at my first day of university straight from the airport as my wife and I returned from our honeymoon, and I have been studying ever since. She worked for years in a job she wanted to leave, unable to take a risk on a new career so that I could pursue mine. Her unshakable loyalty is increasingly rare, and whilst we have shared incredible moments together - we have stood side by side through difficult times too. She simply could not mean more to me.

It would of course be impossible to individually thank all of the friends, family, fellow students and conference delegates to whom I am indebted without exceeding the permitted word count - but they should not think they are forgotten. I have needed them all. Having been amply warned that the successful completion of a PhD thesis tests an individuals' resilience and stamina just as much as their intelligence and knowledge, I was prepared for the struggle that I will readily admit this journey has been. I did however underestimate the unflinching faith required from those around me whilst they patiently waited for me to reach the finish line, all the time knowing I would. They never doubted me, even when I questioned myself - and on reflection, it has been their belief that ensured retreat or surrender was never an option. I hope they all know the part they have played throughout my life, and that if they are proud of me - it is they who deserve the credit.

ABSTRACT

Many people have turned to see someone behind them due to a ‘sense’ they were being watched. Others have ‘inexplicably’ felt as though they were the focus of others’ attention, despite there being no conventional means via which this could be detected (Sheldrake, 2003). The most popular and enduring of the theoretical explanations for these events is that extrasensory awareness was evolutionarily advantageous, and therefore may have developed during an era in which danger was ever-present with survival depending on such capabilities (Sheldrake, 2005).

The term ‘extrasensory perception’ is often abbreviated to ESP and was adopted by Rhine (1934), a Duke University psychologist who employed it to refer to the claimed reception of information gained via the mind, rather than the recognized physical senses. Such abilities include telepathy, intuition, psychometry, clairvoyance, clairaudience, and their associated trans-temporal operations as retrocognition or precognition. Such phenomena are also often referred to as a ‘sixth sense’ or ‘second sight’.

Evidence supporting the existence of extrasensory surveillance detection would have implications beyond purely scientific interest, yet the phenomena remains under-researched and may benefit from a fresh approach. The research conducted as part of the current thesis aimed to examine not only the possible existence of covert surveillance detection - but also which psychosocial and neurological factors may predict this ability.

Research concerning an individual’s ability to detect attention which they could not be aware of via conventional senses has previously been restricted to the psychic staring effect, also known as scopaesthesia - a phenomenon in which people respond via non-conventional means to being the subject of another person’s gaze (Sheldrake 2003). However, this new investigation furthered the research by incorporating the previously uninvestigated sense of being *listened to* as well as seen. The existence of these abilities was gauged during a series of experiments and was measured via a) the accuracy of participants’ self-reports of being watched or listened to, b) psychophysiological reactions determined by electrodermal activity (EDA) which measures the electrical conductance of the participant’s skin to indicate a response, and c) differences in their behaviour during surveillance.

Self-reports (Colwell et al., 2000; Peterson, 1978; Sheldrake, 2000), behavioural differences (Chen, 1937; Cottrell et al., 1968; Dashiell, 1935; Platt et al., 1967; Travis, 1925; Triplett, 1898), and EDA (Colwell et al,. 2000; Peterson, 1978; Sheldrake, 2000; Williams, 1983) have all provided significant evidence of extrasensory detection in previous research, however they had never been combined in a single study; doing so provided the opportunity to cross reference the results, and to directly compare these methodologies.

This original and unique fusion of neuroscientific, parapsychological, forensic, anomalistic, and psychosocial factors represented an essential and progressive step in understanding possible covert surveillance detection, and its psychosocial and neurological predictors such as schizotypy and temporal lobe lability - and it produced findings both expected and surprising. Through a series of studies which were adapted and improved upon based on the results of the experiments which preceded them, the researcher was able to uncover not only which methods of covert surveillance detection were the most effective, but also the circumstances under which they were most sensitive.

Analysis of the resulting data revealed individuals to be able to self-report surveillance they could not be aware of via conventional senses as the literature would suggest (Sheldrake, 2003), however the importance of addressing participant expectation and the reporting bias associated with this was made clear. A major finding of the research however was the discovery that peoples' behaviour could be significantly altered by covert surveillance, as results demonstrated that participants' decision-making ability was affected by whether they were being watched and/or listened to during a cognitive task.

Perhaps the most surprising element of the collection of experiments though was that through the evolving methodology, it was revealed that the stress involved in being tested may be a necessary element for effective research to be conducted in this area. Indeed, by creating and adding stress to experiments in which it was previously absent, the researcher was able to capture positive results via participants' EDA even though this physiological measurement had been shown to be an ineffective measurement of surveillance detection when the participants were relaxed. When similar results were found following a field experiment based on the laboratory research, the researcher developed a theory that stress or threat is an essential element which should be included in future research related to this topic, as well as considered in real-world environments.

CONTENTS

TABLES

FIGURES

INTRODUCTION

THESIS OVERVIEW AND THE IMPORTANCE OF STUDYING SURVEILLANCE DETECTION VIA NONCONVENTIONAL MEANS.

The following thesis incorporates a body of experimental research conducted to investigate the neuroscientific, parapsychological, forensic, anomalistic, and psychosocial factors and implications related to the possibility of surveillance detection via nonconventional means. The aim is to discern what constitutes this so-called extrasensory detection, whether there is any credible data to support such a notion, and which psychological factors may predict this ability.

This thesis will offer an original and scientific approach to investigating the popular and widespread belief that it is possible for an individual to know if they are being watched or listened to. Research has found that a large proportion of individuals report having turned to see someone behind them due to a ‘sense’ they were being watched. Others have ‘inexplicably’ become aware of a conversation involving them, despite it being inaudible. This is often depicted in various novels and movies, but these occurrences do not only take place in fictional settings and are often reported in everyday life. Furthermore, this ability does not seem to be rare in frequency or restricted to just a few ‘gifted’ individuals. Both European and North American surveys have revealed that between 70% and 97% of those questioned stated they had personally experienced such instances (Braud et al., 1990; Cottrell et al., 1996; Sheldrake, 1994).

Theoretical explanations put forward by those who believe it may be possible to somehow detect attention through psychical means include the popular idea that such extrasensory awareness is evolutionarily advantageous (Sheldrake 2003), and therefore may have developed during an era in which danger was ever-present with survival depending on such capabilities. However, should the ability to detect being stared at be a genuine human attribute, it seems likely that evolution by natural selection will have occurred, and so raises questions regarding *how* may it have evolved (Sheldrake, 1999) and does it still benefit those who can access such capabilities?

This extrasensory detection is therefore considered to be beneficial in promoting survival rather than being an abstract and random phenomenon, and is related to ‘psi’ - the theoretical capacity to influence matter or acquire information via non-conventional means which are as yet unexplainable through scientific understanding. Psi subsumes the term psychic ability, which includes psychokinesis (PK) and extrasensory perception (ESP). The former refers to the psychic influence of matter, whilst the latter refers to the psychic acquisition of information and incorporates clairvoyance, precognition and telepathy (Thalboune, 2004).

Should ESP exist, it would have significant implications for covert surveillance as awareness may alter the behaviour of those under investigation. Considering the current and seemingly ever-increasing pervasiveness of surveillance, be it covert or otherwise, the forensic applications of its unconventional detection seem obvious and necessary - yet are under-researched. However, that is not to say they are *un*researched. Sheldrake (2003) coined the term scopaesthesia to refer to the phenomenon in which people are able to detect being the subject of another person's gaze via non-conventional means. He also found through an extensive series of interviews that the sense of being stared at is not unfamiliar among surveillance personnel, police officers, and soldiers. Indeed, not only did army snipers report that they believed certain people could sense being observed, but also that this occurred even over a substantial distance when the target was being viewed through telescopic sights.

On virtually any other subject, anecdotal testimony from such highly trained and trusted professionals would be considered reliable evidence; however, the idea of sensing another person's attention is met by some with a scepticism (Marks & Colwell, 2000; Wiseman, 2010) that only the most robust experimental results could challenge. In fact, many consider the notion of scopaesthesia to be unexplainable by conventional science. For this reason, the current research not only examined the possible existence of this apparently 'psychic' detection - but also which psychosocial and neurological factors may predict this ability, thus extending the current body of work in this area to seek a possible explanation for extrasensory detection, as well as its possible existence in relation to surveillance.

Surveillance is defined by the Oxford English Dictionary (2010) as 'close observation, especially of a suspected spy or criminal' which makes clear the obvious importance of the subject. Whilst observation is sense system neutral and can indicate that the target is being watched or listened to, it is commonly thought of as visually based. This is often the case with hidden cameras, CCTV, satellite observation, or the physical following of an individual - however phone tapping, bugging, computer surveillance, and hidden microphones are increasingly topical and relevant in the modern world. Therefore, the investigation incorporated the sense of being *heard* as well as seen, which is termed acoustasthesia (Friday & Luke, 2014) to further the existing psi research in terms of both scope and methodology.

Psi refers to the unknown factor in experiences that cannot be accounted for by currently understood mechanisms, either physical or biological. The subject is usually researched within the investigative fields of parapsychology which tests for the existence of such phenomena, anomalous psychology which seeks to explain it via conventional science, or psychical research which investigates mesmeric, psychical and spiritualist events. However, the current research aims to add a psychosocial and neurological perspective to this potentially important brain function which, if demonstrated to exist would challenge the basic assumptions and paradigms within psychology and all related sciences. As such, for the purpose of this thesis it was accepted as a working hypothesis that extrasensory perception may be possible, even if understanding of this ability may be limited, and that it may not necessarily exist at all - hence a parapsychological approach.

The thesis therefore takes the form of a critique of the literature and a review of current extrasensory detection understanding as well as the psychology, neurology and psychosocial aspects proposed as being related to it. This fusion of methodologies, measurements and factors offered the chance to investigate extrasensory ability in previously unattempted ways, and to examine which psychological variables may predict individuals' capabilities and belief.

Outline of Thesis

The structure for the remainder of the thesis is as follows. Chapter One is comprised of a literature review of the study of psi, remote observation experiments, and research related to extrasensory detection. It describes the previous investigative methods, findings, obstacles and issues encountered so far. It also describes the predictors of surveillance detection chosen to be examined in the current thesis, and outlines the rationale for their selection.

Chapter Two describes the creation, testing and development of the first empirical laboratory-based experiment referred to throughout the thesis as Study One, which was designed to discover whether an individual can detect being watched and/or listened to via non-conventional means measured by self-reports, physiological response, or changes in behaviour. An investigation and exploration was made of the possible relationship between individuals' ability to detect covert surveillance and several psychosocial and neurological predictors; namely anomalous belief and experience, temporal lobe lability, and schizotypy. The effectiveness of self-reports as a measure of surveillance detection were assessed by asking participants to read words aloud and indicate when they believed periods of surveillance were occurring. Automatic reaction to surveillance was then examined by monitoring physiological response via EDA as this may indicate an unconscious ability to detect surveillance, and lastly, in the third protocol the same participants were asked to undertake a short cognitive task known as the Stroop Test (Ridley, 1935) under four conditions (neither being watched or listened to, being watched, being listened to, and being watched and listened to) to examine differences in their response times. The inclusion of this latter measure gave the experimenter chance to test whether the 'social facilitation effect' or 'reactivity' which is described as peoples' tendency to behave and perform differently when alone to when they are in the presence of others (Griffin & Kent, 1998) would make a reliable means of covert surveillance detection. For all three protocols, the participants knew that these periods of being watched and/or listened to would occur, but did not know the order or duration. This chapter also features the results of Study One, and outlines the process of re-coding the data for accurate and meaningful analysis. It concludes with a discussion and explanation of the findings in which the methodology used and how this could be amended is considered, and directions for future research are suggested.

In Chapter Three, the flaws associated with the first study are considered further, and the testing of Study Two is described with the important methodological changes justified via reference to the experimenter's findings up until that point, as well as the existing literature. The analysis of the second study's results are reported on, with particular attention drawn to how the methodological changes between the first and second studies, or other factors may account for any difference in outcomes throughout the discussion section which delves deeper into comparisons, similarities and important differences between how the two studies were conducted. The limitations of laboratory based experiments and how relatable they are to real-life situations are also considered.

With the ecological validity of laboratory-based experiments discussed, the design, testing and development of a field-based study referred to in the thesis as Study Three is described in detail in Chapter Four with reference to related literature. The results of this field study follow, and are analysed in the discussion section which interrogates the possible reasons for the outcomes, the problems encountered, and possible directions for future research of this kind.

With two laboratory studies and a field study designed, conducted and analysed; the data from all three are compared and contrasted in Chapter Five, and the findings of the current research project as body of work as a whole is considered. What the results mean for the fields of forensic psychology, social psychology and parapsychology are discussed, whilst implications for covert surveillance in the real world as well as possible future directions for the subject are suggested. The considerable list of articles, journals, book and conferences without which this thesis would not have been possible are listed in Chapter Six, whilst the materials and SPSS output necessary for the current study can be found in Chapter Seven.

1 SURVEILLANCE DETECTION VIA NON-CONVENTIONAL MEANS: A LITERATURE REVIEW

1.1 What Is Covert Surveillance Detection and Why Should It Be Studied?

1.1.1 Extrasensory Detection and How It Relates to Psi.

The experience of human existence includes phenomena which cannot entirely be explained by conventional science, existing sensory perception, logic or simple delusions. The term *psi* is employed to describe the unexplained factor in anomalous experiences and is defined as "a correspondence between the cognitive or physiological activity of an organism in its external environment that is anomalous to generally accepted basic limiting principles of nature (Palmer, 1986, p. 139). It seems that only when all conventional methods of influence or information transfer are excluded, can anomalous events be attributed to psi. As there is no agreed upon theory, the term psi does not imply how psychic events might occur (Irwin, 1999), and it was originally proposed as an atheoretical neutral term that did not presuppose a mechanism for psychic events (Thouless, 1942). Despite this, psi is often referred to as if it is a process, however it is simply a construct for that which is not yet understood (Broughton, 1991).

Words such as 'paranormal' or 'psychic' are used to define psi related phenomena which mean 'beyond normal' and 'of the mind' respectively (Miller, 2006). These expressions are surprising considering apparently paranormal events are not only common, but the majority of people believe in their existence (Ross & Joshi, 1992). Indeed, research suggests that up to three-quarters of the general population claim to have actually experienced such phenomena (Broughton, 1991).

Various surveys have also found that the majority of individuals not only believe in the paranormal, but that this often has a noticeable impact on their daily lives (Watt & Tierney, 2014). During recent years, parapsychology research has begun to appear in mainstream psychology journals (Bem, 2011; Storm et al., 2010a, 2010b) which have featured detailed reviews of the evidence both for and against these phenomena (Cardeña et al., 2015; May & Marwaha, 2015). There has, however been no integration of the current data and theories.

Various psychologists have been actively producing supportive research, whilst others have dedicated themselves to criticizing the methodology and analysis of the experiments within the field. Both viewpoints can be explained by their expertise related to the relevant areas of memory, perception, belief, and conscious and non-conscious processes - however, it is likely that many of these psychologists lack the required knowledge of the subject (Cardeña, 2018). Indeed, an informed psi sceptic wrote, "Most psychologists could reasonably be described as uninformed sceptics - a minority could reasonably be described as prejudiced bigots - where the paranormal is concerned" (French, 2001, p. 7).

It is easy to forget that at their origin, psychology and parapsychology were not clearly defined disciplines that were distinct from one another. In fact, renowned foundational figures of the former often supported the latter (Cardeña, 2015a; Sommer, 2013) including Sigmund Freud, Bekhterev, Binet, Luria, Ramón y Cajal, Fechner, Hans Berger (the first person to test the electroencephalogram on humans), as well as American Psychological Association presidents Gardner Murphy and William James. In recent years, faculty from well-respected universities such as Princeton, Harvard, and Stanford have all endorsed the continued research on psi (Cardeña, 2014).

Parapsychology has also made important contributions to methodology integrated into mainstream psychology. They include, but are not limited to the introduction of use of randomisation and masking procedures (Hacking, 1988), meta-analysis as long ago as 1940 (Gupta & Agrawal, 2012), studying eyewitness reports, hallucinations, and hypnotic and dissociative phenomena (Hövelmann, 2015). It is surely therefore an important area of scientific interest to fully understand these events regardless of their validity, because even if a so-called 'normal' or conventional explanation is possible - individuals' willingness to believe them to be extrasensory is an important aspect of the human experience. Despite this, such ideas seem to be both divisive and controversial, thus provoking emotional and extreme reactions from believers and sceptics alike (Broughton, 1991; Freedman, 2005) and so impartial, unbiased and open-minded research is essential.

Psi related incidents and belief are the topic of study for parapsychologists, also known as psychical researchers or anomalous psychologists, who investigate supposedly paranormal processes and phenomena, "which in one or more respects exceed the limits of what is deemed physically possible" (Thalbourne, 2004, p. 206). Such sensations include, among others, precognition whereby an individual appears to be able to receive anomalous information from the future, telepathy which involves the direct transfer of information between two living organisms, or telekinesis defined as direct mental influence over the physical world. Extrasensory Perception (ESP) is the main focus of this thesis however, and the term refers to a human being's apparent ability to acquire information without using their ordinary senses or logical inference. This phenomenon can be attributed to anything from vividly 'seeing', dreaming or envisioning an event, to having a vague intuition, or acquiring information that a person is not consciously aware of - but affects his or her behaviour regardless (Broughton, 1991, p. 33). Radin (2004) criticised this definition however for neglecting to mention that ESP can affect the physiology of the recipient organism. Regardless of the definition, such events seemingly demonstrate a direct interaction between the external environment and the mind, which does not appear to be mediated by accepted sensory or motor means. Because of this, such experiences are deemed theoretically impossible according to mainstream scientific worldview, and therefore beyond the accepted range of human capability (Thalbourne, 2003).

The majority of modern research in the area of extrasensory perception and psi in general is experimental and conducted within a laboratory environment (Broughton, 1991) as researchers are then able to address two of the main issues faced when testing for psi ability; namely whether conventional explanations for the phenomenon have been ruled out via rigorous experimental controls, and whether a supposed psi effect is a statistically significant occurrence. The former is vital as if there is even the slightest chance that the results may have been caused by anything other than psi, the psi hypothesis can be neither supported nor rejected.

This area of research has been repeatedly investigated, but according to Wiseman (2010) the body of work has so far yielded results so inconsistent that the differences may be attributable to reasons other than the existence or non-existence of the phenomena itself. In fact, for more than a century scientists have investigated the possibility of extrasensory ability (Edge et al., 1986), and perhaps more than any other area of psychology, these studies have been divisive and controversial. Proponents claim that the research supports the existence of such phenomena (Bem & Honorton, 1994; Utts, 1991), whilst sceptics argue that the research is flawed with methodological and statistical inadequacies (Alcock, 2003; Hyman, 1994).

1.1.2 What Is Covert Surveillance Detection?

In order for an individual's sensation or 'feeling' that they are being watched to be truly influenced by psi, awareness of remote observation must be achieved without ordinary sensory channels by which the awareness of such surveillance could be conventionally transmitted. In other words, the individual who is being observed, must be unable to see or hear that the observer is watching them.

Belief in remote observation detection among the general public is high, reported to vary between 70% to 97% depending on the population in question (Braud et al., 1993a: Coover, 1913; Sheldrake, 2003; Thalbourne & Evans, 1992). Whilst the high prevalence of anecdotal and spontaneous reports of apparent remote observation detection cannot be considered as evidence that such experiences are anything more than coincidental reactions, the result of confirmation bias, or behaviour driven by normal sensory means - they do warrant experimental investigation to determine whether such sensations are a genuine psi phenomenon. As such, researchers have attempted to test whether remote observation can be detected under experimental conditions that eliminate sensory leakage (information acquired by a person via conventional means) to eradicate ordinary and logical means of detection.

Such studies have been ongoing for over a century, and by reviewing them much can be learnt regarding the issues associated with the methodology, as well as the various explanations offered for the findings so far. Research into this topic began with studies in which the observer and the observed were together in the same room. Following criticism based on the possibility that sensory leakage may account for positive results, these relatively simple experiments progressed to include methodology which ensured that the observer and observed were separated sensorially and physically as they were positioned in different rooms with video links alone making observation possible. These experiments can broadly be divided into two types; those which required conscious 'guess' measures, and those which utilised unconscious psychophysiological measures such as electrodermal activity (EDA) to detect remote observation detection (Wiseman & Schlitz, 1997).

1.1.3 What Does Covert Surveillance Detection Feel Like?

Papers which detail the observed participants' feeling of being remotely stared at are relatively few. First observed in the psychological literature and described as an "uncanny… unpleasant tension" by Titchener (1898, p895), the sensation is anecdotally reported to be a negative experience. Titchener suggested that this feeling was similar to the nervousness many people would naturally experience from knowing someone is behind them. Coover (1913) collated descriptions from his participants and divided them into either visual images or kinaesthetic sensation categories. Visual imagery included unprompted images of the observer staring at them directly, whilst the kinaesthetic sensations associated with being the target of another persons' gaze included restlessness, connection or closeness to the observer, discomfort, being criticised, and wanting to turn around. Only one of the studies reviewed mentioned the feeling of not being observed, which was described as feeling alone.

Other researchers have found that the feeling of being watched evokes negative emotions and sensations. Sheldrake (2003a) amassed a considerable collection of remote observation detection reports and found that even within the spontaneous anecdotal accounts, the overwhelming sensation is articulated as uneasiness, followed by restlessness. It should be noted that positive feelings have not been associated with spontaneous cases of remote observation detection. According to a survey conducted by Thalbourne and Evans (1992), 78% of parapsychology students in Australia experienced an emotional reaction to being stared at such as anxiety, fear, or anger when under remote observation. The researchers also found that 56% of the same sample experienced physical reactions including a pounding heart or blushing.

Literature concerning observation in general seems to suggest that whilst most people find being stared at unsettling, in certain cases this can create positive feelings if the bond between the observer and the observed is close - this is known as affiliation (Schachter, 1959). Anderson (2012) suggested that perhaps such feelings could be initiated via remote observation detection and offered the term 'remote affiliation' to describe it. Anderson also considered that the negative experiences reported such as uneasiness and nervousness may be due to remote observation detection originating from an evolutionary process which allowed our ancestors to be aware of a predators' gaze. This would seem to make sense when considering Sheldrake's (2003a) discovery that individuals who would benefit from this ability such as Special Forces operatives, snipers, security guards and police officers who report instances in which the phenomenon is apparent to them, as their target's survival would to some extent depend on their ability to use such senses.

1.1.4 What Should We Call This Phenomenon?

In scientific terms, the feeling of being watched is sometimes referred to as 'remote observation'. In this instance, 'remote' is used to denote "inaccessible through ordinary means" (Braud, 2003, p. 27), however the observer does not necessarily need to be physically distant, but they must be remote in the sense that conventional sensory pathways which may inform the observed that he or she is being watched must be impossible. The phenomenon has also been referred to as 'remote staring' (Wiseman & Schlitz, 1997), the 'unseen gaze' (Braud et al., 1990; 1993a), and the 'remote staring effect' (Sheldrake, 2005a) - however all of these terms refer to the same supposedly extrasensory ability which is defined as a psi-mediated awareness of another person's stare. As such, the 'sense of being stared at' is an often-used term but is both too vague and cumbersome to be the subject of detailed and robust experimentation, however a clearer definition of the sensation is surprisingly difficult to agree on.

Indeed, the idea of remote observation has caused debate among various respected researchers. Sheldrake (2005) highlighted the possible interpretations and noted that without clarification, 'the sense of being stared at' could include the less controversial idea that animals are aware when others are watching them. The term is often assumed to refer to a humans' ability to detect being stared at from behind beyond the visual range - however for clarity it should be added that the phenomenon in question is indeed related to detection that is extrasensory. In other words, such awareness is not the result of detection via conventional senses.

Interestingly, the use of the word 'sense' has even been questioned for its appropriateness. Braud (2003) raised the point that for the person being stared at, the detection may be accompanied by a feeling and physiological changes which do indeed justify a term such as 'sense' being employed to describe the phenomenon - however in many instances the detection is indicated by behavioural reactions. An often-reported physical response to being stared at is to turn around, despite the person having no conscious awareness of the staring taking place, although it should be noted that the staring might not necessarily be the cause.

Braud is not alone as Atkinson (2005) has also questioned 'sense' as a useful term - stating that if such a sense contained just a simple message detailing nothing more than the information that someone is staring, the ability would be somewhat primitive and therefore unlikely to be the result of evolution over millions of years. Whilst this may appear to make logical sense, the vast majority of experiments to date focus on dichotomous 'yes' or 'no' responses when investigating such phenomena. The idea that there could be more to the ability to detect attention than merely knowing whether one is being stared at or not is supported by peoples' claims that they are aware of the direction from which a stare is emanating, which area of their body is being stared at, and the intentions associated with the stare.

Baker (2005) supported 'remote staring detection' as an accurate description, but Sheldrake (2005) suggested that whilst this is entirely appropriate for experiments conducted using CCTV, it applies less to close range staring detection. Sheldrake therefore prefers 'non-visual staring detection' but proposes that the scientific term 'scopaesthesia' (2003) is less cumbersome and provides an improved definition. For succinctness and clarity, scopaesthesia will be used throughout this thesis to refer to a human's possible ability to detect being stared at without the use of conventional senses, with or without the use of CCTV.

The word scopaesthesia comes the Greek verb *skopein,* which means 'to look at', and *aesthesis*, which means 'sensation'. The latter can also be found in the more common words *anaesthesia*, which means 'no sensation', and *kinaesthesia* which means 'sensation of movement'. Scopaesthesia appears to be the best possible terms as it implies 'feeling' or 'sensation', and so incorporates the term 'sense', but in scientific terminology aesthesia also implies detection - thus apparently answering all the objections detailed above.

1.1.5 How and Why Might Scopaesthesia Be Possible?

Sheldrake has suggested that if it is really possible to detect remote observation, then theoretically it must have been subject to evolution by natural selection (1999). Interestingly, the sense of being stared at does not appear to be restricted to a phenomenon that may occur strictly between humans. This potentially supports the idea that the ability to detect being watched can be explained by evolution and could have developed during a time when humans were both the hunter and the hunted (Sheldrake, 2003) and early detection was essential to survival. If credence is to be given to this popular theory, then it logically follows that the ability would extend both to our prey, and the creatures that would have preyed on us.

Some modern-day hunters and wildlife photographers are unfaltering in their belief that animals are able to detect a human's gaze even when conventional detection through sight and sound is impossible. Furthermore, they are convinced the ability remains even if the animal is being watched through telescopic sights (Sheldrake, 2003a). This is also reported to work in reverse, with hunters and photographers sensing the glare of wild animals (Corbett, 1986; Sheldrake, 2003a). Indeed 34% of adults and 41% of children questioned in an Ohio State University survey reported that they have sensed animals staring at them, with around half the respondents indicating the reverse - that they believed an animal which could not see them had reacted to their gaze (Cottrell et al., 1996). It should of course be noted though that whilst this level of reported behaviour seems powerful evidence for scopaesthesia, it does indicate that roughly the same amount of respondents had noticed no such behaviour.

Sheldrake also proposed a radical theory of perception that offers an insight into how remote staring may be possible (1994). He posits that contrary to commonly held assumptions, people do not see images of the world around them inside their brains. Instead, Sheldrake suggests that these images may in fact be *outside* an individuals' brain. According to this theory, vision may be a two-way process that involves an inward movement of light and an outward projection of mental images. This hypothetical process has interesting implications. Firstly, if peoples' minds can reach out and effectively touch whatever they look at, this suggests that staring may directly affect the object of attention - thus to use the example of remote staring, the observed may be able to literally feel the observer staring at them. This sensation of being stared at apparently feels like there is pressure from skin, tendon, muscle, and joint in or around the region of the observed person's neck according to the reports of those who have supposedly experienced the phenomenon (Sheldrake, 2003).

Historically, there have been two basic theories of vision which largely fell into the category of extramission (emissions from the eyes to the object), and intromission (emissions from the object to the eyes). Sheldrake (2005) addressed this directly when responding to commentators such as Blackmore and Koch who claim that 'scientific' or 'normal' theories of vision must, by definition be the result of intromission - one of science's most venerable theories. According to Sheldrake who explained that its origins date back to the early seventeenth century, various quantum mechanical approaches seem to open up the possibility of a two-way process with regards to vision. Sheldrake also noted that most researchers commenting on his work also clearly stated that if scopaesthesia is a real phenomenon, it would be incompatible with these conventional scientific normal theories. However, Clarke (2004) argued that these are not real alternatives as quantum entanglement may not only play an essential role in vision, but is also an important aspect of conscious perception. He stated that "if the qualitative aspect of perception (the so-called qualia) are produced by quantum entanglement between the states of the brain and the states of perceived objects, then the supports of conscious loci are not just the brain, but the whole of perceived space. In other words, 'I' am spread out over the universe by virtue of my connectivity with other beings' (2002, p 177).

Velmans (2007) described how conventional science says that the colour, shape, location and visual features of an object individuals perceive are just surface representations of the object's appearance which has been constructed by the brain's visual systems. This can be demonstrated by neurological syndromes in which certain features of the visual system have been damaged. For instance, without achromatopsia (a condition characterised by a partial or total absence of colour vision), the visual world would appear to be entirely black, white and various shades of grey. Similarly, in the case of other syndromes, an inability to see form, movement, or depth in space exists. In fact, the surface representations constructed by a fully functioning visual system are not complete without representations of those surfaces. Velmans explained that theses surface appearances are completely different to the descriptions of an object's deeper structure and the physical space in which they are embedded given by physics, (relativity theory and quantum mechanics for example). So, although individuals usually treat the phenomenal object as though it really is the 'physical object', what they experience is nevertheless how that object looks to that particular person, and not how it actually is.

Although humans tend to think of the three dimensional space in which a perceived object is embedded as 'physical space' - this too is how that space looks to them, which is the phenomenal space, rather than the actual space itself. It therefore follows from this that whilst perceived objects are 'physical' in the sense that really do exist and do have an appearance, they are also 'psychological' in another sense as the way that they appear depends not only on the objects itself - but also on the way its appearance is constructed within ones visual systems. So, individuals do not have an experience of an object 'in their mind' or 'within their brain' in addition to that object as they perceive it out in the world. Rather, these phenomenal objects comprise what is experienced - and in terms of the object's phenomenology, an object as perceived and ones experience of that object are the same thing. To explain his point, Velmans (2007) gave an example; he described how when looking at the words on a page, the words that one sees is their only experience of those words. He used this to argue that the naïve realist view is what individuals perceive 'out in space' is the actual object itself, and the individuals also have an additional, veridical experience of the object in their mind or brain is incorrect in two ways - it is not consistent with either third person science, nor the first person experience meaning there can be no real difference in the subjective vs. objective status of the phenomenology of objects or stimulus.

Even Velmans' (2007) perceptual projection theory is 'scientific' and 'normal' from the point of view that he suggests perceptual projection is non-physical, and so does not violate the intromission theory. But, this raises the question of what exactly is perceptual projection? Velmans (2018) explained this as being a common, psychological effect that is readily observable and produced by individuals' preconscious mental processes. Perceptual projection may appear mysterious, but it has been extensively researched - albeit under different names by psychological science. Examples include how peoples' experience of depth can be explained as a construction of the mind/brain arising from cues arranged on two-dimensional surfaces in stereoscopic pictures as is the case with 3D cinema, holograms, and virtual reality (Velmans, 1990). There are also the underlying processes such as the perception of location and distance in space, or information in the light that contributes to depth perception (Hershenson, 1998), as well as the neural structures which support it (Goodale & Milner, 2004). There are also instances where depth perception can break down (Robertson, 2004), and studies to show how the judged metrics of phenomenal space can relate to physical measurements of space (Lehar, 2003) and how these can relate to neural state space. As neural state space is by definition in the brain, and that phenomenal state space is located predominantly outside the brain - an understanding of the relationship between neural state space and phenomenal state space would also provide a topology of perceptual projection. (Velmans, 2009, p. 162).

A greater understanding of perceptual projection would also offer a more unified knowledge of wide ranging phenomena experienced to have both location and extension. Examples include lucid dreams, virtual reality, eidetic imagery, hallucinations, the construction of a body image, as well as the normal perception of events in three-dimensional space. Acceptance of perceptual projection as a normal effect when perceptual processes form representations of real-world events makes it easier to understand what is happening in artificial situations. Velmans (2018) gives the example of three-dimensional virtual worlds which can be understood to arise from artificially stimulating the same projective processes that create the normal phenomenal world. Similarly, hallucinations can be understood to arise from mental models which erroneously project information that originates internally, rather than externally. Likewise, the projections, transferences and counter-transferences that result from therapeutic interactions may be understood as similar internal vs. external confusions where one's personal feelings, thoughts or memories are bound into one's projected experience of another individual. Because the processes that achieve 'binding' and 'projection operate pre-consciously, one can literally experience another's to manifest the qualities and traits which are actually one's own. (Velmans, 2009, p. 163)

Whilst studies of such varied phenomena all contribute to the understanding of the processes and cues that contribute to projective effects, they cannot completely explain just how proximal neural causes within the ones brain support experienced events that appear to be outside ones brain. To explain this, an explanatory model is needed, and no adequate explanatory model exists currently. Virtual reality and projected holograms provide enticing analogies (Velmans, 2009) to the extent that these virtual realities generate perceived three-dimensional worlds and provide an existence proof for projective psychological effects. In doing so, they provide creative ways to study how information in the light can provide cues on which projective processes may operate. Projection holograms provide an analogy for how information encoded on a two-dimensional surface can be perceived as a three-dimensional object in front of that surface if viewed from the appropriate perspective. There is however little convincing evidence that a literal 'neural projection hologram' exists in the brain. Given the lack of any adequate explanation, the emergence of scientific models which respect the three-dimensional phenomenology of conscious experiences, whilst providing a guide to the projective neural processes that support them would be an interesting and exciting development (Velmans, 2018).

Both Lehar (2003) and Gray (2004) proposed that the phenomenal world in its entirety is a version of 'virtual reality' which is located on the inside of an individuals' brain, and that these projections also require a 'vehicle' or 'ground' similar to a holographic projection process. This is effectively a field model, although it is not regarded as an electromagnetic field in a literal sense and so the 'ground' of the projection remains obscure. If, as Velmans (2007) insists, such a projection process is non-physical - it is difficult for one to conceive how it could be related to the brains physical processes, to the electromagnetic field of light, or to scopaesthesia.

Lehar (2003) believes it to be more ‘scientific’ to think of the perceptual projection being located inside the brain, although Sheldrake (2005) points out that this logically leads to the unlikely conclusion that when one observes the sky, their skulls must be beyond the sky they are perceiving. Conversely, Velmans (2007) theorises that the projection would be located outside the head, just as people perceive it to be. He is however, cautious not to inadvertently imply extramission theories as he is not suggesting that the projection occurs through a person’s eyes.

As people experience the world by looking through their eyes rather than the tops of their heads though, the perceptual projection hypothesis makes more sense if it did occur through the eyes. Clarke (2004) suggested a helpful way of interpreting the standard theory of vision and proposed that the perceiver and the object they are perceiving may already be linked together by an electromagnetic interaction which is typically understood to be one way. In this sense they are not individual self-contained systems, rather they should be considered as a single entity. Clarke posits that a dual-aspect view of this system should be adopted to enable consciousness to be associated with brain activity as well as the electromagnetic field. In this way, part of this consciousness can be associated with the position of the object being perceived, which therefore has to be *outside* the brain.

Clarke’s suggestions are useful, however Sheldrake (2005) suggests a dual aspect of the electromagnetic field does not entirely explain remote staring, and offers reflection and refraction as examples for two reasons; firstly, he points out that images are not an aspect of the electromagnetic field, but instead are split off it into virtual space. Secondly, he highlights that consciousness may not be necessarily connected to the electromagnetic field - but may instead be selectively connected to it. Sheldrake explained that when people look at a reflection in a window, they can choose to focus their attention on that reflection or look past it and through the window instead. In other words, what the person sees depends on their attention, even though theoretically the same electromagnetic field is linking their eyes to what they are seeing. Whilst it could be argued that this is simply the shifting focus of attention, Sheldrake (2005) argues that this would appear to demonstrate that perceptions must be more than just lights entering a persons’ eyes.

How auditory and visual elements could relate to a sense or feeling of being watched or listened to were hinted at through the work of Martinez et al. (2012) who posited that spatial frequency is a fundamental visual feature which is coded in the primary visual cortex, and that its relevance relates to the perception of objects, textures, scenes, and hierarchical structures (Shulman et al., 1986), in addition to eye movements and the direction of attention (Özgen et al., 2006). Temporal amplitude-modulation (AM) rate is an essential auditory feature which is coded in primary auditory cortex (Liang et al., 2002) - it is relevant for the perception of auditory objects, speech and scenes (Shannon et al., 1995). Spatial frequency and temporal AM rate therefore form the fundamental building blocks of auditory and visual perception. Research results have suggested that crossmodal interactions are common across the primary sensory cortices (Yau et al., 2009) and that some underlying neural associations develop via consistent multisensory experiences such as audio-visually perceiving gender, speech, and objects (Schwartz et al., 2004).

Martinez et al. (2012) demonstrated that individuals consistently and completely (rather than relatively) match their specific auditory AM rates to their specific visual spatial frequencies. This crossmodal mapping therefore allows amplitude-modulated sounds to direct attention to and modulate awareness of specific visual spatial frequencies. Results also showed that based on physical spatial frequency, crossmodal association is approximately linear and generalizes to tactile pulses. This suggests that such an association is developed through multisensory experiences during the manual exploration of surfaces.

Results demonstrated that observers reliably matched a specific visual spatial frequency to a specific auditory AM rate - this remained the case even when each observer viewed just one visual spatial frequency. Such a finding suggests that the underlying crossmodal mapping is not relative, but is absolute. The function relating visual spatial frequency to auditory AM rate was shown to be approximately linear, and significantly less compressive than the function which relates spatial frequency to location on a line scale. This suggests a unique perceptual relationship between auditory AM rate and visual spatial frequency that is separate from an abstract magnitude representation.

Such a crossmodal association would allow sounds to influence visual processing of spatial frequency, so even when AM sounds were task irrelevant and were ignored, and observers reported no awareness of the auditory-visual association - they remained able to more quickly discriminate the direction of a phase shift if it occurred on the crossmodally congruent Gabor (a linear filter used for texture analysis). This finding suggested that an AM sound can guide attention to the corresponding spatial frequency via increasing its salience. Such an interpretation is consistent with binocular rivalry results found by Martinez et al. (2012), as despite observers ignoring the sounds and reporting no awareness of the auditory-visual association, an AM sound was shown to increase the proportion of perceptual dominance of the congruent Gabor - it would therefore appear that the AM sound boosted visual signals (Kim et al., 2006).

Previous studies demonstrated that crossmodal interactions were strong if both auditory and visual stimuli were attended, but were reduced when just one modality was attended or when attention was redirected by a more demanding concurrent task (Degerman et al., 2007). Additionally, if the auditory-visual association was defined only by rhythmic synchronization, a sound was found to influence binocular rivalry only when it was attended (Van Ee et al., 2009). It is therefore possible that any effect of AM sounds upon visual attention and awareness found in the experiments conducted by Martinez et al. (2012) could have been even stronger if their observers been made aware of the crossmodal associations and told to attend to the AM sounds. Regardless, the absence of awareness of the associations was essential in the experiments for addressing concerns regarding the issue of response bias.

In answer to the question of what could be the neural substrate of the association between auditory AM rate and visual spatial frequency - Martinez et al. (2012) offered the following explanation. It could be the neurons in an individuals' primary auditory cortex (A1) tuned to AM rate (Liang et al., 2002) and those in their primary visual cortex (V1) tuned to spatial frequency (Geisler & Albrecht, 1997) interacting (Ghazanfar & Schroeder, 2006). However, the experimenters deemed it unlikely that the perceptual relationship they demonstrated directly involves V1 neurons as they are tuned to retinal spatial frequency, whilst the perceptual association that was obtained depended on physical spatial frequency.

As physical spatial frequency is thought to be perceptually more salient than retinal spatial frequency, individuals typically discriminate (Burbeck, 1987) and rapidly remember (Bennett & Cortese, 1996) physical spatial frequencies that convey distance-invariant information useful for object recognition, rather than their retinal equivalents (Sowden & Schyns, 2006). Therefore, the coding of retinal spatial frequency can be rapidly transformed into the coding of physical spatial frequency within mid or high-level visual areas exhibiting relatively size-invariant responses to visual patterns (Eger et al., 2008). Martinez et al's. (2012) findings may therefore indicate a unique relationship between AM-rate-tuned neurons in A1 and physical spatial-frequency-tuned neurons in intermediate or high-level visual areas.

This crossmodal neural association may form because many surfaces are approximately periodically corrugated and because individuals may move their hands at relatively constant speeds while exploring surfaces manually, the rate of AM sounds people hear and the rate of the pulses they feel while sliding their hand over a surface may both correlate experientially with physical visual spatial frequency. Martinez et al. (2012) posit that this manual-exploration hypothesis, auditory AM rate and tactile pulse rate are similarly associated with physical (instead of retinal) spatial frequency.

In summary, Martinez et al. (2012) demonstrated a fundamental association between the temporal processing of auditory AM rate and the spatial processing of visual frequency. This link appears to be absolute and unique as it is approximately linear, and is likely to be distinct from an abstract magnitude representation, and unaltered even when observers determine the auditory match to one solitary visual spatial frequency. This relationship allows an AM sound to direct attention to and increase awareness of the corresponding visual spatial frequency, and does so by crossmodally modulating the visual signals in a spatial-frequency-specific way. Such a linear dependence on physical spatial frequency and the generalization to tactile pulses could indicate this link is formed via the multisensory experience of manually exploring surfaces, and could begin to help explain how the sense of being stared at, and a possible sense of being listened to may operate.

1.1.6 Who is Capable of Scopaesthesia?

There appears to be no particular type of person who reports the feeling of being stared at, with the majority admitting to personally experiencing such sensations (Braud et al., 1990; Sheldrake, 1994; Cottrell et al., 1996). So perhaps extrasensory surveillance detection is not only to be found in fictional tales by the novelists such as Dostoyevsky, Tolstoy, and Doyle (Poortman, 1959). However, Sheldrake (2005) found potentially important gender bias when he carried out related surveys in Britain, Sweden and the United States and discovered that 81% of women had reported the feeling of being stared at, whereas only 74% of men had experienced this sensation.

Taken at face value, these figures seem to suggest that women are better at experiencing remote observation, but it could be that there are differences in the way they translate or remember the experience.

It is also worth noting that the experience of scopaesthesia tends to be reported most often between strangers in public places (Sheldrake, 2003a). To be able to confidently make such claims however, gender differences in peoples' willingness to report and admit to such experiences should also be a subject of investigation as this may account for what initially appears to be a gender-based difference in detection capabilities. Furthermore, it is important to note that the same survey investigated the 'feeling' of being stared at which could occur even when the experiencer is not under observation, and does not consider the accuracy of this sensation. As such, no claims can be made regarding peoples' actual ability to detect remote observation when they really are the subject of another persons' gaze, just that they reported a sense or feeling associated with being stared at.

When it comes to the so-called 'ability' to make other people turn around by simply staring at them, the gender gap seems to widen with 88% of women reportedly able to detect observation, as opposed to 71% of men (Sheldrake, 2003a). Again, such claims require tightly controlled empirical experimental research to support them as this data was obtained via self-reported responses to a survey, however such statistics hint that when it comes to extrasensory detection, perhaps ability may be linked to necessity and how important it is to the individual to know that they are under surveillance. This thinking relies on the notion that women generally feel more vulnerable than men, and so this idea, along with the situations and circumstances under which these reports are made should be considered. This theory has further implications also, as if necessity and importance directly affect extrasensory ability - this will have repercussions for experiments where there are no high-stakes or threat involved. It could therefore be extrapolated that finding no apparent effect of scopaesthesia in laboratory-based experiments in which there are no consequences to a participants' failure to detect surveillance would not necessarily mean that scopaesthesia is not a genuine phenomenon in situations involving actual and imminent danger.

1.1.7 Why does Extrasensory Detection Matter to Science?

The very idea of scopaesthesia conflicts with this conventional 'intromission' theory of vision published by Kepler (1571-1630) in 1604. Kepler was best known for his discoveries in astronomy, but his explanation of vision suggested that light enters the eyes, and that this is a one-way transaction meaning that vision is inside a person's head, and not in the outer world. Thought of as one of modern sciences great triumphs, 'intromission' theory of the retinal image appeared to resolve a two-thousand-year old debate, but it also raised a problem that even Kepler himself admitted he could not solve; it remains an unanswered question as discussed in section 1.1.5.

Kepler's theory describes how images form on the eye's retina, but failed to explain sight itself. People see the outside world not upside down and double, but as a single image the right way up. In other words, people do not see a pair of tiny inverted images of the external world on their retinas. To deal with this troublesome issue, Kepler simply excluded it from optics (Lindberg, 1981; Winer et al., 2002). As such, the intromission theory of vision effectively left 'seeing' unexplained. By relegating the process to the brain's interior, it has become the psychologists' problem to solve.

Bach-y-Rita (1972) stated, "we see with the brain, not the eyes." Cortical plasticity of the brain enables individuals to interpret imagery information on visual terms (Bach-y-Rita, 1995), although perception of the image relies on other non-psychobiological factors, such as memory, learning, and cultural and social factors (Bach-y-Rita et al., 2003). Bach-y-Rita (1995) shows from the psychobiological perspective that the formation of imagery in visual terms occurs when perceptual levels of the brain interpret the spatially coded neural activity, which is then modified and augmented by non-synaptic and other brain plasticity mechanisms.

Farah et al.'s (1989) findings highlights the role of psychobiological influence in imagery formation, as their findings showed that mental imagery involves the modality-specific visual cortex, due to the slow, late positivity, maximal at the occipital and posterior temporal regions of the scalp when participants were exposed to visual imagery. However, the role of memory may also contribute to the formation of mental imagery, as previous knowledge of sensory activity may help create mental imagery in the absence of sensory input (Frith & Raymond, 1997).

The theory of hearing is also very mysterious, and scientists still have much to learn. What is known however is that sounds are invisible vibrations travelling through the air known as sound waves. When these waves reach fibres with a resonant frequency, they release bursts of energy, which in turn move tiny hair cells located in the organ of corti - a structure that stretches from one side of the cochlea to the other (Peng et al., 2011). The cochlea is only capable of sending raw data though, which are transmitted as complex electrical impulse patterns. The person's brain then takes this input and deciphers it like a central computer. Whilst the concept itself is relatively straight forward, the structures and process are incredibly complex. As scientists continue to make advancements, they discover new hearing elements every year - and so research on the brain's role in listening and hearing is potentially of value beyond the current study.

1.1.8 Why Does Scopaesthesia Matter to Current Affairs?

The term 'surveillance detection' will be used throughout the current thesis. This is often thought of as something confined to the secret service and national security, which the everyday person would only encounter when reading books or watching films set in the world of spies, espionage, and terrorism. In reality though, surveillance detection is reported every day, by ordinary individuals living typical lives. So common are these experiences that they are often disregarded as meaningless despite the fact that researching psi phenomena has wide reaching implications for understanding human behaviour, the limits of our abilities, and the nature of consciousness and reality.

This topic is becoming increasingly relevant though, be it through global concerns such as the necessity to detect terrorist attacks before they occur, or members of the public who are worried their everyday activities may be observed or listened to. Increasingly, the individuals who feel as though they are being tracked or are under surveillance may not just be paranoid, as monitoring really is often taking place. For example, public transit audio surveillance may be monitoring the conversations of those onboard American buses as authorities have installed surveillance equipment in the hope that recording and monitoring the audio may yield important information and that a major security issue may be detected among the mundane exchanges of passengers (Kille & Maximino, 2014). In fact, an international survey of public transport organisations (UITP, 2015) revealed transport operators worldwide are investing in security technology, with the majority introducing an increased use of real-time surveillance despite 97% of survey respondents having security cameras installed already. This will include both closed circuit TV (CCTV) cameras and audio recording devices - and so everything from the materials passengers read, to what they say to each other will be seen and heard, as will their mobile phone conversations.

Even without transport surveillance, people's mobile phones can be remotely tapped - and not necessarily by covert government spying organisations or highly skilled computer hackers. Such technology is readily available, and requires no more than software easily downloaded from the Internet. The owner of the mobile device does not even need to be on the phone to be listened to, as the programs allow the phone's microphone function to be engaged remotely even when not in use. Similar software is being developed by law enforcement agencies that can be implanted wirelessly, then delivered and installed via text message. This means that anyone who 'feels' like they are being listened to, may well be.

Even in the office environment, a global network of computers named ECHELON is capable of monitoring workers' e-mails, phone records and web surfing on behalf of several world governments. The network's creators have also devised a way to monitor every piece of paper that goes through a laser jet printer with a microscopic code identifying the specific printer that the paper came from. It takes specialist equipment such as blue lights and magnifiers to read it, but it can be used to give away behaviour and habits - and companies such as Dell, Xerox, Lexmark, Canon and others have begun installing this surveillance technology. Even personal laptops have the capacity to be hacked, and the webcam within it used to remotely watch the user.

The monitoring of people and their behaviours has an inarguable value when used to thwart crime and terrorism, but manufacturing giants and superstores are also very interested in the ability to keep shoppers under surveillance via RFID Chips - microscopic radio frequency identification devices that can be implanted into almost anything. The RFID chips purpose is to allow retailers to track which products have been sold. This technology is meant to become inactive once the product has left the store as they are affectively 'turned off' by the cashier, but should they forget, the shopper is effectively under surveillance all the while they are in possession of the item.

Similarly, the Terrorism Information Awareness Program (Weinberger, 2008) tracks peoples' electronic transactions daily searching for merging patterns to help the government determine individuals who are possible threats, and so individual's actions and behaviours are constantly monitored. If it is possible to be aware of this via extrasensory means, evidence of this phenomena would have considerable implications for people diagnosed with paranoia - a thought process associated with anxiety or fear to the point of delusion (American Psychiatric Association, 2013), as the feeling that they are the object of others' attention may not actually be irrational.

1.1.9 Scepticism - Why the General Public may Dismiss Psi.

Stanford (1982) and Broughton (1992) have both argued that people may possess psi abilities, but use them unintentionally and unconsciously to automatically predict and affect events and outcomes. If this were the case, those individuals who competently use their extrasensory abilities and are therefore "psi-effective" (Broughton, 1991, p. 352) would do so unknowingly, and so would not necessarily report psi related experiences, nor necessarily believe in their existence. Additionally, individuals with a reduced tendency to observe patterns in nature may attribute genuinely anomalous experiences to random chance. According to Sheldrake (2005), people dismiss the possibility that they are able to detect the attention of others via extrasensory means simply because this would be classified as 'paranormal', despite most people having actually experienced the phenomena themselves. The subject is therefore considered taboo and rejected as a superstition by many, despite personal experience to the contrary, and irrespective of research findings.

1.1.10 Scepticism - Why Some Scientists Dismiss Psi.

Whilst it is reported that a significant majority of people can tell when somebody is staring at them (Sheldrake, 1994), these claims are contested on scientific and statistical grounds. Marks and Colwell (2000) dispute Sheldrake's research due to the sequences used not being properly randomized, and insist that when they are - peoples' accuracy in detecting remote staring is no better than chance. These authors state that Sheldrake proposes ideas that could possibly exist, but, in all probability, do not. They also highlight how conducting controlled scientific experiments is challenging, and that this is especially true of psi related research as results are difficult to support if effects are inconsistent and may be attributed to unintended variables.

Whilst Marks and Colwell (2000) also take issue with Sheldrake's willingness to rely on the participation of amateurs and non-scientists to conduct research on his behalf as they are free to explore new avenues of research, they also concede that at times researchers do seem reluctant to accept new ideas and paradigms. Indeed, when psi related research is left entirely to a few professional scientists, this can produce its own problems (Marks, 2000; Marks & Kammann, 1980) such as experimenter bias, the file drawer problem, concerns regarding reputation, repetition of previous mistakes, and location issues. However, issues with correct randomisation, double-blind controls, proper prevention of cueing, independent judges, and proper statistical procedures are also often cited as reasons not to take evidence for psi seriously.

An example of this critique is illustrated in Wiseman's (2010) *Skeptical Inquirer* article where he argued how Parapsychologists supposedly nullify null results. He described how, in his opinion parapsychologists tend to view positive results as support for the existence of psi, whilst adopting strategies to make certain that null results do not count as evidence against the existence of psi. Specifically, Wiseman detailed how experimenters 'cherry pick' new procedures meaning that positive findings in parapsychology have emerged from a collection of non-significant studies, yet are more likely to be presented at conferences or published in journals than non-significant results - however, Carter (2010) pointed out that Wiseman is unable to cite evidence for his claims. Additionally, Wiseman's article discussed how parapsychologists may explain away unsuccessful attempted replications and create excuses for not accepting replications which show evidence opposing the existence of psi. Similarly, Wiseman claims that parapsychologists decide retrospectively to only analyse data that shows evidence of psi to ensure that meta-analyses offer their preferred conclusions.

1.1.11 Defence Against the Sceptics.

When considering the many criticisms of psi related research, many are able to defend it using the same scientific and statistical theories as the sceptics (see section 1.1.10). A typical question posed by the sceptics of psi is to simply ask why not *all* studies replicate positive results if the reported phenomena are real. However it has been argued (Barrett, 2015; Lewontin, 1994) that one should actually expect some replications to fail if the small effect sizes found in 'successful' experiments are considered - especially if potential variability sources such as psychological and parapsychological experimenter effects (Palmer & Millar, 2015) are also taken into account. Harvard professor Robert Rosenthal (1990) shared this opinion, stating that "given the levels of statistical power at which we normally operate, we have no right to expect the proportion of significant results that we typically do expect, even if in nature there is a very real and very important effect" (p. 16; see also Utts, 1991).

Further argument against those sceptical of positive psi results can also be found in the explanation of why parapsychology effect sizes are typically small, and in how they compare to other areas of psychology. Unfortunately, psi experiments are subject to ethical considerations, and as such involve impersonal stimuli which is not comparable to real-life circumstances. Such stimuli are often of little consequence, and so are in stark contrast with those often reported to be related to psi phenomena such as the unexpected passing of a close relative, or a near-death experience (Cardeña, 2018). Additionally, psi appears to be more reliably manifested by a relatively small number of individuals, so the effect sizes are most likely to be the average of larger effects of selected participants and smaller or null effects of their counterparts (Harris & Rosenthal, 1988).

In relation to other areas of psychology, Richard et al. (2003) analysed more than 25,000 social psychology experiments and reported an average effect size of 0.21 - similar to the effect size found in some of the parapsychology meta-analyses conducted by Cardeña (2018). It should also be noted that the effect size of some psi protocols is more than just comparable with those of the clinically recommended uses of propranolol or aspirin to prevent heart conditions (Spencer, 1995; Utts, 1991), as it some cases they are in fact larger - and so would be classified as 'evidence-based' if the criteria of clinical practice were applied (Haidich, 2010).

It would therefore appear that psi effects cannot be replicated 'on demand' (Bem et al., 2001; Bem et al., 2015; Cardeña, 2018; Gilbert et al., 2016), and this can be put in perspective if the 'Many Labs' project is considered. This extensive work incorporated 36 independent laboratories in an attempt to replicate 16 psychology experiments - all of which were published in top journals. Just 34% of these replicated studies fell within the confidence intervals of the original version (Open Science Collaboration, 2015). Despite this evidence, proponents of parapsychology also find themselves defending against sceptics' claims that significant psi effects are the results of low-quality experiments, however Cardeña's (2018) meta-analyses of such experiments controlled for the quality of such studies, and significant effects still remained.

Indeed, the allegation of sub-standard experimentation and selective publication (known as the file-drawer problem) is the source of much frustration for parapsychologists as they are often the victim of such practices. For example, psi critics were eager to publish their failed replications of Bem's studies, but the same could not be said of experiments in their database which supported his findings (Ritchie et al., 2012). Similarly, it is claimed that studies which would support psi have not been submitted for publication as they were conducted by skeptics (Sheldrake, 2015).

The file-drawer problem was also addressed in a direct rebuttal to Wiseman's *Skeptical Inquirer* article (2010), when Carter argued that a "heads I win, tails you lose" strategy is often adopted by sceptics to discredit parapsychologists' findings. In response to Wiseman's claims that "parapsychologists frequently create and test new experimental procedures in an attempt to produce laboratory evidence for psi. Most of these studies do not yield significant results. However, they are either never published or are quietly forgotten even if they make it into a journal or conference proceedings." (p. 37), Carter responded with a question. He asked exactly how Wiseman knows that most of these studies do not yield significant results if they were unpublished, as there appeared to be no evidence for his accusation. Carter also highlighted that due to the controversial and divisive nature of anomalous phenomena, parapsychologists were in fact among the first to tackle the issue of, for instance, the file drawer problem.

This attempt to address the problem of the selective publishing of positive results at the expense of unreported non-significant results is an important step in the history of parapsychological research. Indeed, the wider psychology community would do well to follow this example as the current 'replication crisis' (Schooler, 2014) is by no means unique to specialist fields. In 1975 the Parapsychological Association adopted a policy unique among the sciences which opposed the withholding of non-significant data. In addition to this, in 1980 the sceptical psychologist Susan Blackmore conducted a survey of parapsychologists to investigate a possible bias towards reporting successful results; she concluded that there were none. Carter also stated that he believed Wiseman (2010) to be hypocritical when stating "any failure to replicate [the original effect] can be attributed to the procedural modifications rather than to the nonexistence of psi. Perhaps the most far-reaching version of this 'get out of a null effect free' card involves an appeal to the experimenter effect, wherein any negative findings are attributed to the psi inhibitory nature of the researchers running the study" (p. 37).

Carter (2010) pointed out that one of the most well documented studies demonstrating this so-called experimenter effect involved none other than Wiseman himself (Wiseman & Schlitz, 1997) when he and Marilyn Schlitz ran the same study, using the same equipment, in the same location - but obtained entirely different results. The subject of this research was remote observation, and so is detailed later in this thesis (section 1.2.4), but the experiment perfectly demonstrated the effect to which Wiseman refers to in his critique. His results were no different from chance, while Schlitz's experiments produced statistically significant data.

On the subject of meta-analyses and retrospective data selection, Wiseman (2010) wrote that "Parapsychologists have tended to view positive results as supportive of the psi hypothesis while ensuring that null results don't count as evidence against it" (p. 36). However, Carter again challenged this and stated that this is confusing absence of evidence with evidence of absence. In other words, failure to find an effect in a solitary experiment cannot count as evidence that the effect does not exist, as individual studies may not show positive results for various reasons including improper execution and issues with sample size. Again, Carter accused Wiseman of using similar techniques when conducting a meta-analysis of the results from thirty Ganzfeld psi experiments (Milton & Wiseman, 1999) when the researchers failed to consider sample size.

Carter concluded his rebuttal by highlighting that psychologists and researchers keen to dismiss or ensure they never find positive results for anomalous phenomena are able to achieve the results they wish to present by ignoring commonly accepted statistical techniques, and arbitrarily excluding data that opposes their desired results. He encouraged researchers to remember that many scientific controversies have strong ideological components, and for this reason poor, but well disguised science can sometimes be conducted.

Whilst no consensus exists regarding psi's existence, and the many researchers who claim its apparent effects can be explained by conventional science may well be absolutely correct, it is also worth bearing in mind physicist Max Planck's words (1950, p 33-34) who as one of quantum mechanics founding fathers remarked "A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it."

Whilst a psi sceptic, Cornell professor in statistics Greenhouse (1991) supported Planck's statement by declaring "parapsychologists should not be held to a different standard of evidence to support their findings than other scientists" (p. 388). A sentiment echoed by Deming (2016) who also concluded that there should not be two different types of scientific evidence, and criticised the sceptics arguments saying they are intended to "suppress innovation and maintain orthodoxy" (p. 1319).

Such robust defence can often leave sceptics resorting to the argument that exceptional claims require exceptional evidence, although this is problematic as many phenomena that we do not consider in any way to be 'exceptional' such as electricity were historically thought of as extraordinary, or even impossible. A requirement for the same 'exceptional evidence' demanded of parapsychology might have prevented such phenomena from ever being accepted. This is further complicated by the term 'exceptional' being difficult to define, open to interpretation, and meaning different things to different people.

Science's history demonstrates that previous certainties can be undone by new discoveries, findings and theories. After carefully investigating the field, some researchers such as the distinguished University of California neuroscientist James H. Fallon (2015) have become convinced that psi research exhibits "methodological excellence" (p. 12) and promising results. Such open-mindedness would be in the spirit of William James and his fellow founders of psychology, who believed in a comprehensive and open discipline (Cardeña et al., 2017).

1.2 Researching Surveillance Detection by Extrasensory Means

1.2.1 What Have We Learnt So Far About Scopaesthesia.

Whilst the covert surveillance theme of this study means that one of the experimenters' primary aims is to investigate whether it is possible for an individual to be aware that they are being watched or listened to via non-conventional means, existing research in this area is entirely dedicated to scopaesthesia and is therefore restricted to the sensation of being under visual surveillance. The audible element is an entirely original additional, and its inclusion will therefore be investigated. For this reason, the methodology of the current study will be predominantly built on previous research which has focused exclusively on peoples' ability to detect covert visual observation. These experiments have been many and varied, and the range of methods, variables and theories employed will be detailed here from the subject's origins over a century ago, to recent and sophisticated laboratory-based experiments.

1.2.2 Classic Scopaesthesia Studies.

The first investigations into remote observation required the observer and the observed participant to be in the same room. These experiments adhered to a basic procedure in which the individual trying to detect observation would be facing away from the person observing. The observer would then either look intently at the back of the observed participant's head in the observation trials, or look away entirely in the non-observation trials whilst the observed would consciously attempt to guess which trial was taking place. Tests of this kind have been conducted with participants of all ages resulting in tens of thousands of trials (Sheldrake, 2003). Results are often positive, with approximately 55% of the 'guesses' correct - this is statistically significant when taken across so many trials.

The first documented study to employ this now classic setup was conducted in 1898 when Titchener discovered that some of his students believed that they could sense being stared at. However, he rejected the notion of telepathy, and his subsequent laboratory experiments conducted with participants who insisted they could sense other people staring at them yielded ‘invariably’ negative results. Unfortunately, Titchener made this statement without providing the results or reasons for his conclusion. It is therefore unfortunately impossible to assess his claim.

Similarly, Coover (1913) conducted 1,000 trials using 10 participants who believed in their own ability to detect remote observation. Observation and non-observation condition sequencing was determined via the roll of a dice to ensure randomisation, and the observer signalled the start and finish of every 20-second trial by tapping their pencil on a desk. Upon completion of the experiments Coover found his participants to be correct in whether they were being stared at just 50.2% of the time, which he suggested was an ‘astonishing approximation’ of pure chance. However, when subsequently re-analysed by Sheldrake (1998) with emphasis placed on analysing the observation trials in isolation, the hit rate in Coover’s study was found to be significantly above chance at 53.3%. This may indicate that participants were able to detect being watched, but were unable to detect *not* being watched.

Seeking to improve Coover’s methodology, in 1959 Poortman sought to recreate the experiment - but increased the duration of the trials to investigate whether covert surveillance detection is an instant or quick reaction to being stared at. Poortman believed that he could remotely detect observation and so was the observed participant in his own study, and he recruited a female observer alleged to be exceptional in influencing other individuals by staring at them. Together the pair took part in 89 trials lasting approximately three minutes each, and all took place in the observer’s house.

It would be interesting to know whether this amount of trials was pre-planned as ‘optional stopping’ (the deliberate ceasing or extension of studies to reach the results desired) is an often-cited explanation for positive outcomes (Wiseman, 2010). The order of conditions for every trial was decided by the observer drawing a playing card from a pre-shuffled pack; the colour of the card determined the condition. Trials were started by the observer knocking on the table at which she was sat, and then spoke out loud to conclude each trial. With 59.55% of the ‘guesses’ correct overall, results were significantly positive.

These experimental results appear to be replicable, with 20 out of 21 of the classic observer and observed participant pairing studies reviewed by Sheldrake (2005) showing positive outcomes, however some researchers have still challenged the idea that people can sense being stared at. Indeed, Sheldrake's experiments have been criticised for failing to use randomised patterns

(Marks & Colwell, 2000), with his claims also branded unfalsifiable (Shermer, 2005). In stark contrast to Sheldrake's positive results, a psychology student project at the University of Amsterdam found that in one out of three experiments they conducted, the results were at chance levels; whilst the participants scored more positively than negatively in the other two, but not with statistical significance (Lobach & Bierman, 2004) despite using the same methodology as Sheldrake.

Sheldrake (1998, 1999, 2000, 2001, 2005b) highlighted that response biases may play a role in the results. For instance, if participants report being *observed* more often than they report being *non-observed* that may mean more 'hits' in the observation condition - but could also help ensure misses in the non-observation condition. Sheldrake noted that this bias to answer positive more than negative does exist (2005b), but that it is not the only explanation for apparent remote observation detection as the corresponding 'misses' in the non-observation condition do not occur. It could be assumed that fraud, cheating and cueing would result in increased 'hits' in observation and non-observation trials alike, and so are unlikely to be the only explanatory factor, but a combination of these with response bias could be.

This collection of studies demonstrates the origins of research into remote observation detection, but they are heavily criticised and suffer from important limitations. Perhaps one of the most crucial flaws with the methodology was the lack of consideration for sensory leakage. Examples include, but are not limited to the sound of the observer shifting positions, changes in their breathing, shadows or reflections hinting as to their position, and differences in the volume and accuracy of the noises intended to start and end the trials which could have influenced differences between observation and non-observation. Indeed, some participants even reported that they deliberately listened for such cues (Coover, 1913). Additionally, the observation condition sections may not have been sufficiently random, and could have perhaps matched the observers' guessing patterns.

As with any experiment, the risk of error, fraud and deception is difficult to eradicate completely and so no single experiment can provide conclusive evidence supporting or refuting claims that the detection of remote observation is a genuine phenomenon - research has therefore continued within the same paradigm. More recent studies have employed similar experimental protocols (e.g., Sheldrake, 1998, 1999, 2000, 2001; Schwartz & Russek, 1999), with Sheldrake accumulating an immense amount of data on the topic via workshops, school classroom experiments and volunteer third-party researchers (Sheldrake, 1998, 1999, 2000, 2001, 2003, 2005b).

Trial durations were typically between 10 and 20 seconds long, and their order was either randomised ahead of the trials or based on real-time coin tosses or rolls of a dice. Methods of starting the trials varied between the observer or an experimenter speaking, or by a mechanised clicker. Interestingly, many of the experiments involved the participants receiving trial-by-trial feedback regarding whether the 'guesses' were correct. Sheldrake (2003a) has reported that such feedback can have a dramatic effect on the accuracy of 'guesses' and a later section of this thesis (1.2.7) is dedicated to this. There are however risks involved with providing feedback - specifically, that providing this to the observed can create artefacts in the process as they intentionally or unintentionally learn patterns in non-random trail sequences. Another potential problem with providing feedback is that the observed participant may begin to associate cues derived from sensory leakage with certain observation conditions. Unintended sensory leakage of any kind would be problematic in any experimental design, but giving trial-by-trial feedback could exaggerate the issue.

1.2.3 Does Personality Affect Scopaesthesia?

Lee et al. (2002) investigated whether personality was related to how people may conceptualise being observed, and devised their own questionnaire to measure 'intuition' and 'social awareness' traits. Participants with high intuition scores reported that they felt watched in public places, whilst those with high social awareness scores did so when they felt "alone, isolated or vulnerable" (Lee et al., 2002, p 408). Such findings may indicate that people experience feeling watched in environments that cause them unease, which concurs with the unsettling feelings reported in spontaneous cases (Coover, 1913; Sheldrake, 2003a; Thalbourne & Evans, 1992; Titchener, 1898). It is important to note that researchers should not confuse their participants' feeling uneasy with the sensation of being observed (Anderson, 2012).

1.2.4 Experimenter Bias in Scopaesthesia Studies.

Whilst remote staring has been the subject of experimentation on both sides of the debate, researchers often have a tendency to collaborate with like-minded colleagues whose beliefs regarding the existence of extrasensory ability are the same or similar to their own. Indeed, observational theories suggest that psi experiments could actually exploit the indeterminacy of a system, and that this may become biased due to the intention of the observer (Millar, 2015), or as Stapp (2017) expressed it, by "relevant conditions that include the experienced emotions of biological agents" (p. 106).

The idea that a researchers' intentions may affect the results of an experiment makes Schlitz and Wiseman's (1997) work an important addition to the literature, and it was later supported by Schwartz and Russek's (1999) experiment in which they investigated the idea that imagined observation could test the effect of the observer's intention to observe (see section 1.2.8). Importantly, the 1997 study by Schlitz and Wiseman also offered a valuable contribution to the idea that joint sceptic-proponent collaborations offer the potential to resolve this area of debate and disagreement (Hyman & Honorton, 1986). Schlitz, a proponent of psychic ability, and Wiseman, a well-known sceptic conducted a study on remote staring as part of a series of collaborative research projects and found possible evidence of the experimenter effect - a phenomena for which there is considerable data to suggest that experimenters' attitudes have the potential to influence experimental outcomes in the direction of their own expectations (Rosenthal, 1976). It should be noted however that the alternative explanation is that the difference in the experimenters' own psi ability may account for the opposing results.

Wiseman and Schlitz's (1997) experiment contributed important literature related to the effect of remote observation on electrodermal activity (EDA) in which the electrical conductance of some participants' skin increased when they were under observation, thus suggesting a possible physiological reaction to being stared at. Physiological reaction as a measurement of scopaesthesia is discussed in detail in section 1.4.1, however these researchers' studies also served as an excellent example of the experimenter effect, also known as experimenter bias. The study was created to examine the effects of a sceptic experimenter who had found non-significant results in their own remote observation research (Wiseman & Smith, 1994; Wiseman et al., 1995), with a believer in psi who had reported significant results from their remote observation experiments (Schlitz & LaBerge, 1997). To conclude that any difference in results were due to the experimenter effect, all of the trials had to be run identically; consequently all took place at the University of Hertfordshire in the UK, used the same apparatus, and all of the 32 participants were from the same pool and assigned opportunistically to either Wiseman or Schlitz who greeted them before setting up the EDA equipment.

Randomised and counter-balanced observation sequences had already been prepared, and it was only after the experimenter had left their participant that he or she learned of the observation sequence to avoid cuing or priming. Once the experiment began, the experimenter and the participants being observed were in different rooms separated by 20 metres. Observation was made possible by CCTV, and following a two-minute delay to allow the participant's EDA to settle to a baseline level, the observer either watched the observed participants' image on a monitor, or looked away. Trials were 30 seconds long, and each participant underwent 32 trials. Even with stringent controls for sensory leakage and attention paid to avoid cuing, Schlitz's observation produced a significant increase in the participants' EDA during remote observation trials when compared to non-observation trials. Conversely, when Wiseman was observing there was no significant difference between observation and non-observation in the participants' EDA.

These results suggest that psi proponents will elicit a psychophysiological response from those they observe, whilst a sceptic observer will not - however, it is impossible to know for certain whether the difference in results was due to the experimenters' interactions with the participants, or their own psychic ability. The non-random allocation of participants to observers could have also resulted in a difference between the two groups, which may account for the apparent effect of the experimenter's observing. This risk could have been avoided by comparing the two observers within-subjects, or having a neutral colleague greet the participants.

Schlitz and Wiseman's approach provides an interesting insight into extrasensory detection, research on anomalous phenomenon generally, and why seemingly identical research can elicit opposing evidence as their results were congruous with their differing expectations. As the proponent's experiment yielded significant data whilst the sceptic's did not, this perhaps suggests that the open-mindedness of those conducting the study could be key to the validly of future research, and that methodology will be improved by removing as much of their interaction as possible - or at least by monitoring the experimenters' beliefs for consistent trends.

1.2.5 Does Sample Size and Quality Make a Difference?

It is pointed out that even among the studies conducted by sceptics, their research has produced statistically significant positive results, as well as responses at chance levels (Sheldrake, 2001). It seems reasonable that an increased sample size may provide answers, but such research comes with its own issues. The largest project ever dedicated to the subject of remote staring began in 1995 at the NeMo Science Centre. By 2002, more than 18,700 observed/observer pairs had taken part, with significantly positive results; however, these tests were not supervised and were conducted by members of the public visiting the centre. Consequently, there were no controls over deliberate or accidental cheating, or the opportunity to misreport results - and so their validity must be questioned.

This problem of study quality in terms of controlling for sensory leakage, sample size and randomisation of sequences was addressed by Schmidt et al. (2004) who conducted a meta-analysis of 15 studies with an overall total of 379 sessions and found a non-significant relationship between the quality of the studies and the effect size. The authors also found that overall, there was a significant mean effect size of $d = .13$, ($p = .01$) which suggests the existence of remote observation and that an anomaly related to distant intentions should not dismissed even though there are only hints of an effect. However, they cautiously concluded that whilst the data available indicates an unknown effect, more high-quality research - particularly with large datasets is necessary to determine whether these results merely represent an artefact. The fact that meta-analysis is vulnerable to type I errors should also be considered, as the results may have led the authors to incorrectly reject a true null hypothesis.

1.2.6 Patterns in the Data: Accuracy in the Observed Condition.

According to Sheldrake (1998, 1999, 2000, 2001, 2005b) who has considered the data available in combination, the majority of these remote observation tests give positive scores, and closer analysis reveals a distinctive pattern of data that repeatedly shows up in these results. His observations lead him to believe that peoples' tendency to detect observation more accurately than they can detect non-observation might be due to the stimulus that observation provides and concluded that this is easier to detect than the absence of a stimulus. The pattern's striking consistency across so many independent experiments suggests that the effect is not merely an artefact such as the system of randomization, implicit learning of sequences, or reporting biases.

It seems that the typical success rates in observation trials are usually above chance, but are often at chance levels in the non-observation trials. Put another way, participants seem to know when they are being watched, but not when they are being ignored. It is worth noting that success in the observation trials is not dependent on a gifted minority of particularly sensitive participants, but in fact represents a more general tendency for participants' self-reporting accuracy to improve when they are being stared at, as opposed to when they are not.

This same pattern of success only in staring trials has also appeared in more recent studies (Colwell et al., 2000) when the participants and the observers were separated by a one-way mirror, as well as in a replication (Radin, 2004). This could be argued to suggest that detection would only be possible when the participant is the focus of a person's stare, therefore explaining the positive results under this condition. In contrast, when nobody is staring, the ability is redundant, and the participant is effectively being instructed to detect the *absence* of a stare. This would be a completely unnatural scenario if there is any truth to the theory that the sense of being stared at is evolutionary and was developed to warn of danger; however, this theory is challenged by Schmidt (2001).

Sheldrake has interrogated the data intensely and investigated remote observation detection in depth by breaking the considerable amount of empirical data down to determine whether the correct 'guesses' occurred in the observation trials or the non-observation trials. In doing so, he has uncovered a pattern - namely that the observed participants detected remote observation around 5% above mean chance expectation. This translates to a 'hit' rate of approximately 55%, however non-observation detection rates were not significantly different from mean chance expectation.

1.2.7 Considering the Criticisms of Scopaesthesia Studies.

To determine whether artefacts had potentially exaggerated remote observation detection, Sheldrake (1999) compared previous experiments' results with varying levels of control against sensory leakage. He compared the hit rates of all the studies which signalled the start of the trial with a sound such as a table tap created by the observer, to those initiated by a mechanical clicker as he theorised that if information regarding the observation condition was being conveyed via the tapping volume, it is likely the hit rates would be greater than in the experiments which used the mechanical clickers.

Hit rates were revealed to be almost identical however, and any difference between them was shown to be non-significant. This revelation disputes the notion that in comparable experiments such as Poortman's (1959), the volume of the tapping created an artefactual effect. Radin (2004b) also provided support for Sheldrake's (1998, 1999, 2000, 2001, 2003, 2005) findings when he conducted a pilot study intended to replicate Sheldrake's experimental design, but with improved controls for sensory leakage and randomisation.

Sheldrake (2001) also considered the possibility that observed participants could actually see their observers, by varying whether those taking part in his studies were wearing blindfolds, but again determined that the differences in hit rates were non-significant. The randomisation of the sequences used by Sheldrake were criticised however by Colwell et al. (2000), who proposed that the lack of randomness contained a pattern which participants could learn from the feedback they were given. Interestingly, it seems that if participants are tested repeatedly and receive feedback following each trial, their scores improve significantly (p = 0.003) with practice (Colwell, et al., 2000).

Indeed, giving feedback resulted in an accuracy rate as determined by correct guesses of 90% when eight to nine-year-old children were tested in a German school (Sheldrake, 1998). However, in 2001 Sheldrake tested this by varying the amount of feedback given to his participants, subsequently finding that the difference this made to hit rates was non-significant. The learning of the trial sequence criticism is also contradicted by the significantly positive findings of previous experiments conducted without providing feedback (Sheldrake, 2000, 2001) which demonstrated the same hit rate pattern. Such findings suggest that the ability to detect remote observation may be a genuine phenomenon.

Using a computer program to simulate the pattern of significant hit rates in the observation condition and chance hit rates in the non-observation condition, Lobach and Bierman (2004) made an important contribution to the literature. They found that by combining a guessing strategy based on feedback depending on how many times the observed participant was under observation, and then comparing this to the response bias towards answering ‘observed’ 55% of the time - the same pattern of responses emerged. However, this fails to explain positive results found in the experiments in which feedback was not given. So whilst many of the experiments did not completely control for artefacts, analysis has indicated that positive findings were not necessarily due to sensory leakage, lack of randomisation, implicit learning, or cues.

By Sheldrake’s (2005) own admission, Radin’s (2004) meta-analysis demonstrated that remote staring experiments reveal a highly significant effect which is difficult to explain simply with subliminal cues and artefacts more clearly than his own analysis. Yet still some sceptics argue that the significant effects could be explained by reporting positive outcomes selectively, otherwise known as the ‘file-drawer’ effect, however Radin quantified this possibility by showing that there would have to be between 1,417 and 7,729 missing studies with either null or negative effects, and that these would have to involve between 800,000 and 3,000,000 missing trials to negate the overall positive results. When looked at in these terms, such a situation seems improbable.

Nevertheless, Carpenter (2005) claimed that it is difficult for an unbiased enquirer not to conclude that scopaesthesia is an illusion, however he seemingly selected an arbitrary collection of just five publications to arrive at this conclusion. As three of these publications challenged the notion of remote staring, he decided against its existence. These studies include Sheldrake’s 2005 study, and Coover’s (1913) work which showed positive data relating to subjects and starers in close proximity, yet he classified Coover’s work as a ‘negative’ study. He did the same with the work of Colwell et al. (2000) which demonstrated highly significant positive results in the first experiment. Similarly, Burns (2005) cast doubt on studies involving CCTV on the basis that some employed psychophysiological methods to measure the EDA which were not up-to-date, and so she viewed them with extreme caution. Whereas in the direct staring trials Burns drew attention to possible ‘matching biases’ in the participants’ responses, as well as imperfect randomisation methods. In doing so, a reader could conclude that *all* possible experimental designs are inherently problematic.

Further critique has been put forward by Atkinson (2005) who questioned the possibility of a scopaesthesia signal for two reasons: firstly that it is not possible to measure such a signal independently of a participants’ subjective report - however to accept this argument, one would have to accept the same criticism in relation to more conventional psychological research.

Secondly, Atkinson also stated that such a signal is too far outside the realms of current scientific knowledge to be studied effectively - a statement that does nothing to disprove the signals' possible existence. However, Atkinson did suggest that remote observation trials could be analysed with signal detection theory, which involves comparing the total hit rate and the total false alarm rate.

Whilst he showed that the discriminability index (*d_*) observed in remote staring trials was small when compared with unrelated observations psychology research, the observations to which he referred were made with selected participants, whilst in almost all other remote staring the participants were unselected. Should scopaesthesia be a real phenomenon, *d_* should be higher if especially sensitive participants were selected. Interestingly, Atkinson suggested that the results could be explained by a response bias, particularly if some participants exhibited delusional ideation and schizotypy personality style. In doing so, he made the argument that response bias should be accounted for in future studies.

1.2.8 Does the Body Part Observed Matter? Can Imagined Observation Have an Effect?

To determine whether the part of the body being observed affects peoples' ability to detect observation, Schwartz and Russek (1999) devised an experiment in which the observer would stare at either their counterpart's head, or their back. The observer stared according to a counterbalanced sequence and the trials began with the observer stating they were "ready" aloud, and ended with the observed participant guessing whether it was their head or back that was the subject of attention. These trials were brief, taking only a few seconds with a ten-second interval between them. Results showed that correct answers significantly outweighed incorrect responses. It should be noted though that the paper does not mention whether the counterbalanced sequences were randomised - an important element when trying to avoid participants learning the sequence patterns either deliberately, or subconsciously.

Schwartz and Russek (1999) took their investigation a stage further and conducted another two experiments to compare whether observation of the head or back makes a difference to detection, but this time took the *imagined observation* of the head or lower back into account. The second experiment was the same methodologically as the first, except that the observer kept their eyes closed in the imagined condition and merely envisioned observing the head or back of their counterpart. The researchers postulated that this imagined observation would test the effect of the observer's *intention* to observe only - an idea that could potentially explain the contradicting results found by Schlitz (a proponent of psi) and Wiseman (a sceptic of psi) in their 1997 collaboration which produced opposing outcomes when Schlitz found positive results, whilst Wiseman did not. This was despite their experiment into remote staring being carried out with identical methodology, using the same equipment, at the same location. By taking this idea into account, Schwartz and Russek (1999) planned to eradicate the possible effect of biophysical energy emitted via literal observation. This however did not rule out cues as the imagined observation condition did not control for differences in the observer's breathing, or changes in the volume or tone of their voice. Analysis of the results revealed however that a marginally significant majority of participants were able to guess correctly whether their head or back was the target of observation suggesting that the body part being stared at is irrelevant. Interestingly though, the participants correctly reported actual observation 56% of the time ($p < .003$), whereas 60% of their reports were correct ($p < .001$) in the intended observation condition.

A similar procedure was used in the Schwartz and Russek's (1999) third and final experiment in the series, during which the opportunity for sensory leakage and cueing was reduced by ensuring conventional senses could not account for detection, and the counterbalancing of the observation condition sequences were improved. The researchers also increased the sample size and the number of trials to increase the statistical power of the analysis. With this experiment they found a significant proportion of correct answers with no significant difference between the hit rates for actual and imagined observation, as they were detected with equal accuracy. With the sequence of observation conditions randomised, and opportunities for sensory leakage and cueing controlled for - the results appear to suggest that imagined observation and actual observation are equally detectable. This is particularly interesting when considering Sheldrake's (1994) theory of perception in which he suggested that people do not see images of the world around them inside their brains, but rather that they reside outside an individuals' brain making it a two-way process involving an inward movement of light and an outward projection of mental images.

Such results imply that the observed participants could be detecting the observer's attention or their intention to observe, rather than their actual gaze. Schwartz and Russek (1999) made an important and interesting observation based on these results, noting that they might explain why many sceptical observers do not obtain significant observation detection effects as the observers' intention or attention may be an essential part of their stare being detectable. If true, should an observer focus their intention or attention towards the observed participant when under the non-observation condition, a similar effect would occur if they were under the observation condition. Such an effect would compromise any experiment intended to differentiate the difference between observation and non-observation, and if individuals are affected by the intention or attention of others, then experimenters could be affecting their participants whether they can see them or not. This theory has wider, real-world implications and does not only have ramifications for all experiments in the field, but indeed all studies, jobs or activities which require observation of any kind. Not least, as some employers are keen to monitor their staff productivity or protect against theft, but if the individuals being observed are aware of it happening, they may alter their behaviour accordingly. This idea can also be directly applied to individuals under surveillance for criminal behaviour who may subsequently then avoid the act for which they were being spied on.

To summarise, research into remote observation detection seems to suggest that participants are able to detect when they are under observation more accurately than if they were merely guessing. All of the above studies have one thing in common - they all test whether a person's gaze is detectable by another individual. This methodology makes the experiments comparable to the spontaneous experience model in which there is a direct line of sight, and so it could be argued that situating the pair of participants in the same room makes the experiment more ecologically valid. The price for such validity however, is that it also increases the risk of sensory leakage and unintended cueing. Therefore, it cannot be concluded that remote observation detection is possible through extrasensory means from these studies alone, as the positive results could be explained as artefacts created through poor experimental control. Whilst Sheldrake's (2003) in-depth analysis suggests that cues, sensory leakage, and artefacts of poor randomisation do not account for the positive results - such factors must be eliminated completely before claims that the remote observation detection effect is due to extrasensory abilities can be made. Studies have attempted to control for sensory leakage, but before reviewing them, participants' self-reported 'feeling' of remote observation detection upon which these experiments rely should be considered and explored.

1.3 Studying Scopaesthesia via Conscious Guess Measures

1.3.1 Conscious Guess Measures, but With Physical Separation.

Studies into remote observation detection have continued to rely on conscious guess measures, but have improved on the methodology of the research which preceded them by attempting to control for sensory leakage and unintended cueing. By physically separating the participants and employing more sophisticated techniques of randomising trial conditions, experimenters such as Peterson (1978) found themselves more able to defend positive results against criticism. Peterson conducted four experimental sessions lasting six-minutes each, and nine pairs of participants played the roles of the observer and the observed. The pairs were in different rooms, which were adjacent to each other with a one-way mirror between them - this allowed the observer to be able to see the observed participant; however, the observed participant could not see the observer.

To further reduce risk of noise related cueing, the observed listened to white noise through a set of headphones. There was a training session prior to the trials in which the observed participant was given feedback, so they could learn what being observed ‘feels’ like, however no feedback was given in the actual trials. Once the experiment began, the observer either stared at, or turned away from the observed participant through the one-way-mirror, the order of which was dictated by a randomised schedule. To indicate whether they felt they were being stared at, the observed participant pushed a button. These recorded button presses were later analysed for how accurately they coincided with the periods of observation and the researchers found that there was a significant relationship, with more button pushes during the observation trials. Previous work supporting scopaesthesia was therefore supported, even when the criticisms of sensory leakage and cueing which plagued such experiments were controlled for.

More recently, Sheldrake (2000) provided further evidence that remote observation detection may be a genuine phenomenon when he reported the results of three well-controlled experiments conducted at London schools. These experiments also separated the observer and observed participants adequately - but entire classes of primary school pupils participated at the same time in the classic observed and observer participant pairs. The former wore blindfolds and were positioned outside with their teacher, whilst those who would be observing were inside the school with Sheldrake; all were able to see the participants who were to be observed through the closed windows of the classroom. The two groups were separated by between three and 100 metres.

The observers only received their randomised observation sequence lists after all participants were in position, and they were unique to each group with a view to controlling for pupils simply copying each other's actions. Those observing were told when to start by Sheldrake, who also made a tone sound outside the school for the observed participants to begin the trial. The 20 trials lasted ten seconds each, and a teacher outside told the observed participants when the trial had ended; these pupils subsequently wrote their guesses on paper. No feedback was given to either group, yet the overall results showed that hit rates were significantly above chance.

Sheldrake had this experiment replicated in the USA, Germany, and Canada and found that overall, participants guessed correctly 55.2% of the time in the looking trials and just 50.8% of the time in the not-looking trials. This means that both in London and the other three countries the scores in the looking trials were significantly above chance levels, whilst in the not-looking trials they were not statistically different from the level that would be expected by chance (50%). In combination there were 2544 correct guesses and 2254 incorrect guesses which reaches statistical significance (p = 0.00003).

To encourage the general public to replicate Sheldrake's (1998) research, Colwell et al. (2000) examined the observation condition sequences which Sheldrake had provided. In these experiments, participants who believed that the ability to detect remote observation exists underwent 60 remote observation detection trials for which they were given no feedback. These were then followed by 180 trials for which they were given trial-by-trial feedback. Sheldrake's colleague Shroder was always the observer, and he was in an adjacent room to the observed participants. A one-way mirror ensured that Shroder was able to see the participants who were to be observed - but not vice versa. Whilst the rooms were not sound-proofed, he was careful not to provide noise cues which could be related to observation. The 20-second trials began and ended with the spoken word "trial", and the observed participants were presented with the options 'yes' or 'no' displayed on computer monitors enabling them to register their guesses. The monitors also allowed feedback to be given when appropriate immediately after the observed participant registered their guess.

In line with Sheldrake's hypothesis that when it comes to scopaesthesia, people are able to detect a stimuli such as another's gaze, but not the absence of a stimuli - the results overall were significantly above chance in the observation trials only. The observed participants also guessed correctly at a significantly higher rate than one would expect by chance in the trials for which they were given feedback, but at chance levels in trials for which they received no feedback. It should also be noted that correct guesses did not increase during the experiment's duration. It could be argued that this may have been due to the participants learning what it feels like to be remotely observed during the course of the trials for which they received feedback. Unfortunately, the observation condition sequences were found to be non-random as there were not enough repetitions of the same condition, and so the significant findings could have occurred through participants implicitly learning the pattern of observation conditions.

In consideration of this criticism, Colwell et al. (2000) tested whether inadequate randomisation may have led to artefactual findings via a second experiment which employed genuinely random sequences using random number tables. This follow-up experiment otherwise adhered to the same procedure as their first, but with a different observer. This time Sladen replaced Shroder, and participants underwent 200 trials each with the initial 10% receiving no feedback. This time there was no remote observation detection effect found overall, and participants did not improve over time. This suggested to Colwell et al. (2000) that the positive effect found in their first experiment was due to the participants learning the condition sequences implicitly. However, Anderson (2012) suggested that there were additional differences between the two experiments which could account for the discrepancies in the results. Firstly, the two experiments had different observers, and their own psi ability may have played a role in the outcome. Secondly, the non-feedback condition which preceded the feedback condition in the initial experiment was longer, and so could have effectively been used as practice, thus improving participants' performance in the feedback condition. The sequence of the observation conditions may therefore not be the only reason for the discrepancy between the two sets of results, but the experimenters noted that their research does highlight the importance of truly randomised observation sequences.

1.3.2 Does the Number of Observer/Observed Participants Matter?

Most experiments focusing on the topic of remote observation detection examine the possible effects of one participant staring at another at any one time. However, in 1994 researchers Wiseman and Smith devised a study in which there were approximately three observers and two participants being observed to try and improve the study's ecological validity as there are often other people present, and more than one observer in real-life situations. All participants rotated roles, and so acted as both observers and the observed. The two groups were in adjacent rooms, with a one-way mirror which allowed the observers to see the observed, but not vice versa. The pair of participants who were to be observed were separated by a screen and shared their room with one of the experimenters. Randomised condition sequencing was used, and the experimental design was checked for any issues with sensory leakage such as the observer being heard when moving.

Observed participants underwent six trials lasting twenty seconds each, and rated each one via a seven-point Likert scale for how observed they felt during the period. Upon analysis of the results, the authors found no significant differences between participants' ratings in observation and non-observation conditions. Whilst this study provides support for the notion that genuine remote staring detection does not exist, the null findings could be due to the fact that observed participants were not alone, and therefore may not have felt like they were unobserved, even under the non-observed condition.

1.3.3 Taking Predictor Variables Into Account.

To further the literature, Williams (1983) investigated remote observation detection by including an additional factor - namely belief in psi. Williams recruited fourteen believers in psi, and fourteen disbelievers selected from a large student group who had completed a ten-item belief in psi scale; these participants were to play the role of the observed. Not only were the observer and observed participants in separate rooms for this experiment - the rooms were 60 feet apart. The observer stared at their counterpart on a closed-circuit television link (CCTV) monitor for 12-second periods which were randomly interspersed with 12 seconds of nothing more than a blank screen. Trials began when a tone sounded, and the observed participant pushed a buttoned labelled 'yes' whenever they felt they were being stared at.

When looking for evidence of remote observation detection, Williams found a large effect size (r = .31) and an overall significant positive deviation from chance. This effect size was greater for observed participants who scored high on belief in psi (r = .49), however the results for those who scored low on this scale were at chance levels indicating that they were merely guessing the condition. With this experiment Williams found that 74% of his participants were successful in detecting remote observation via self-reports, even when sensory leakage was eradicated, cueing was controlled for, and condition sequences were properly randomised. In doing so, he supported the hypothesis that detection of remote observation may be psi-mediated and contributed the idea that belief in psi might improve individuals' ability to know whether they are being stared at.

Conversely, Lobach and Bierman (2004) found no evidence for remote observation detection in their series of experiments created to replicate Sheldrake's scopaesthesia studies under conditions of sensory shielding. A total of 188 sessions were conducted resulting in 4784 trials demonstrating over-all hit rates of 50.6% (N = 53), 52.1% (N = 45), and 49.7% (N = 37). The first of these experiments considered the observer's belief in psi as a variable. For this research, the observed participants were stared at during half their trials by an observer who believed in psi, and by a sceptical observer for the other half. As with many previous experiments, the observer and the observed were in adjacent rooms with a one-way mirror in between them, which again allowed the former to see the latter. The observer stared at the left side of the observed participants' face, or turned away and focussed their attention elsewhere - this sequence was dictated by a pre-prepared randomised schedule. Trials lasted 15 seconds and began with a recorded voice which was played in both of the rooms. For their second experiment, Lobach and Bierman (2004) combined traditional conscious guess measures of remote observation detection with psychophysiological measures.

Psychophysiological responses are of particular importance to the current research and so will be reviewed in detail in a later section (1.4.1) dedicated to this measure, however the conscious guess aspect of this second experiment during which the observer stared at their counterpart through a CCTV link from an adjacent room can be reviewed independently. There were 20 observation and non-observation trials which lasted for 30-seconds, and were dictated by a randomised schedule. The observed participant pushed a button to register their guess in a brief pause between each of the trials, however the remote observation detection hit rate did not significantly differ from chance.

In their third experiment, Lobach and Bierman (2004) compared participants who knew each other with those who were unacquainted. Pairs of friends were recruited, and each played the role of the observed and the observer with each other, and then with someone they had never met. In all trials, the observed and observer were in adjacent rooms with a one-way mirror between them allowing the former to see the latter with detection guesses again recorded via the pressing of a button following after each ten-second trial. As with their previous experiment, the hit rate was not significantly different from chance.

1.3.4 Do Conscious Guess Measures Suggest Scopaesthesia is Possible?

The experiments which have employed conscious guessing as a means of measuring the detection of remote observation have included a diverse range of methodologies and have had various types of participants of differing sample sizes, altered the duration of the trials, and used different means of starting them. All have tried to answer the same question however, and so they are conceptual replications for each other.

Replications conducted by varying experimenters should be encouraged as the involvement of many different researchers can reduce the chance of fraud, and also help guard against the risk of positive effects being due to a hidden artefact caused by the procedure or methodology used (Schlitz & Braud, 1997). The conscious guess hit rate of these studies, the results of t-tests comparing these hit rates to mean chance expectation, and the corresponding effect sizes are summarised in Table 1 below.

Table 1

Conscious Guess Hit Rates Compared to Mean Chance Expectation

Experimenter(s)conscious guess hit rates compared to mean chance expectation	No. of observed participants	t	Significant at p<.05	R	Hit rate %
Peterson (1978)	9	2.65	Yes	0.41	54.86
Williams (1983)	28	1.71	Yes	0.31	53.32
Sheldrake (2000)	155	2.14	Yes	0.17	52.6
Colwell at al. (2000) 1st experiment	12	N/Aa	Yes	0.46	N/Aa
Colwell et al. (2000) 2nd experiment	12	N/Aa	No	0.11	N/Aa
Wiseman & Smith (1994)	65	0.93	No	0.12	N/Ab
Lobach & Bierman (2004) 2nd experiment	45	.124	No	0.18	52.1

a Results were reported in four blocks of trials for each experiment. The reported effect sizes are the mean effect size per experiment based on these blocks. No t-test or hit rates were reported for each experiment overall.

b No hit rates were reported due to participants rating how observed they felt on a Likert scale. Results only deviated slightly from chance expectation.

Of the seven comparable studies presented, four are significant and found a higher hit rate than the 50% that would be expected by chance. The significant studies by Peterson (1978), Williams (1983), and Colwell et al. (2000) had large effect sizes of >.3, whilst Sheldrake's (2000) had an effect size of >.15. All controlled well for sensory leakage and cueing, and without any known artefacts, they could appear to suggest the significant hits rates were due to extrasensory ability, although there could have been a randomisation artefact in Colwell et al.'s (2000) initial experiment.

Peterson's (1978) and Williams' (1983) research arguably provided the most robust controls against sensory leakage, but they produced the highest effect sizes, indicating that sensory leakage does not explain significant results. It should also be noted that two of the experiments which produced non-significant results (Colwell et al., 2000; Wiseman & Smith, 1994) were carried out by sceptics, and so the experimenter effect (Rosenthal, 1994) may have influenced the outcome. Whilst not conclusive, it could therefore be claimed that credible evidence for remote observation detection exists.

When reviewing studies on remote observation detection studies prior to 1985 which were measured by conscious guessing, Braud et al. (1990, 1993a) proposed that the data suggested remote observation detection is a genuine phenomenon, although the effect sizes were small. Sheldrake (1999) has argued that effect sizes are small because remote observation detection is likely to be unconscious and stated that "under the artificial conditions of experiments, people are being asked to do consciously what they may usually do unconsciously. Self-consciousness may interfere with their sensitivity" (p. 67). This seems to suggest that remote observation detection is normally non-intentional, and so the experimenter's attempts to measure this phenomenon via unconscious means whilst the participant is not actively trying to 'feel' the presence of staring may improve the chances of detecting an effect.

Braud et al. (1993a, p. 376-377) had also observed that spontaneous reports of observation detection often include "unconscious behavioural and bodily changes rich in physiological content and automatic movements" and noted that they rarely involve higher cognitive functioning. The authors therefore suggested that unconscious measures may be more sensitive, and so may produce a better means of measuring remote observation. Physiological responses would also address the issues caused by cognitive interferences such as response bias and participants' guessing strategies such as when participants alternate their observation and non-observation guesses to maintain an even overall balance (Coover, 1913). The order of these conditions should also be randomised to avoid the participant detecting underlying patterns when trial-by-trial feedback is used, or using a response strategy.

1.4 Studying Scopaesthesia via Psychophysiological Measures

1.4.1 What Is EDA and How Is It Measured?

With section 1.3.4 suggesting that employing physiological responses in scopaesthesia studies may overcome the problems associated with response bias and guessing strategies, it seems prudent to carefully consider its use in any related investigations. When psychophysiological responses have been used to measure a remote observation detection effect, the dependent variable chosen has typically been changes in a participant's electrodermal activity (EDA).

EDA is also known as electrodermal response (EDR), galvanic skin response (GSR), skin conductance response (SCR), psychogalvanic reflex (PGR), and skin conductance level (SCL). However it is referred to, the measure indicates the activity in the sympathetic branch of the autonomic nervous system, which is often referred to as sympathetic nervous system arousal.

A decrease in sympathetic nervous system response is related to feelings of calm, and an increase in activity within the antagonistic branch of the autonomic nervous system is known as the parasympathetic nervous system. Conversely, an increase in sympathetic nervous system activity occurs when the human body responds to stress and is associated with fear, agitation and nervousness. This is indicated by pupil dilation, changes to heart rate, blood pressure, bronchiole dilation, release of adrenaline, and sweating (Silverthorn et al., 2009). The latter of these symptoms is related to EDA and is measured via electrodes placed on the observed participant's non-dominant hand. A small and painless electric current is passed through the electrodes to measure their skin's conductivity, with an increase indicating arousal.

1.4.2 Using Psychophysiological Measures to Detect Scopaesthesia.

Both spontaneous reports and the feelings participants described during empirical experiments into scopaesthesia reviewed in the above sections depict the sensation of remote observation as one of nervousness and uneasiness. It could therefore be extrapolated that such observation would activate the sympathetic nervous system. Many studies which used EDA as means of detecting remote observation have found an effect, but in opposite directions. By reviewing them, possible explanations for this difference can be proposed.

Importantly, the studies which used the ABBA counterbalanced and randomised observation conditions which would most effectively protect against artefacts were Wiseman et al.'s (1995) second experiment, followed by Schlitz and LaBerge's (1997) study, Wiseman and Schlitz's (1997) research, and then Wiseman and Schlitz's subsequent replication (1998). Except for the trials conducted by renowned sceptic Wiseman, all of these studies demonstrated significant effects. As only studies with possible issues with randomisation (Lobach & Bierman, 2004; Wiseman & Smith, 1994) reported borderline effects, poor randomisation cannot adequately be used to explain significant findings.

Likewise, failure to control for sensory leakage or cueing cannot account for significant effects, as all the studies featured methodology which protected against these possible artefacts. Indeed the observer was in a separate non-adjacent room to the participants in all of the studies except one (Lobach & Bierman, 2004). More recently, a replication of Williams' 1983 CCTV study was attempted, but using electrodermal activity (EDA) equipment which monitors skin conductance to detect participants' reaction to being stared at, but a clear-cut effect was not found (Müller et al., 2009).

Calming effects should also be taken into account as research has suggested that remote observation can reduce an individuals' sympathetic nervous system. When considering the calming effects of remote observation as measured by EDA, the observation came predominantly from two individuals, namely Shafer (Braud et al., 1993a, 1993b) and Schlitz (Wiseman & Schlitz, 1999) as Shafer trained the other observers who elicited calming effects. Their remote observation has also produced activating responses however (Braud et al., 1993a; Wiseman & Schlitz, 1997) which were similar in effect size and were large in both cases.

It seems therefore that Shafer (Braud et al., 1993a, 1993b) and Schlitz (Wiseman & Schlitz, 1999) may be capable of achieving a powerful remote observation effect in participants, and that this could be similar in its magnitude whether it is a calming or activating influence. In stark contrast, sceptics such as Wiseman do not appear to exert an influence in either direction when remotely observing, and so it seems the observer themselves can be a factor. This observation is tentative as it is based on a total of just three observers, but such a pattern would at least suggest that researchers should consider decreases in EDA as well as increases when looking for psychophysiological evidence of remote observation detection.

1.4.3 Meta-Analytic Evidence for EDA Measuring Scopaesthesia.

Schlitz and Braud (1997) and Schmidt et al., (2004) have conducted meta-analyses to estimate the combined effect size of these remote observation and psychophysiology studies. Despite the substantial overlap in the research they reviewed, these meta-analyses found different outcomes. The first meta-analysis (Schlitz & Braud, 1997) reviewed 11 studies which encompassed 230 individual sessions - all of which used psychophysiological means to detect remote observation via CCTV between psychically separated individuals. Just five percent of these experiments would be expected to be independently significant by chance, however the researchers found that 65% demonstrated a significant difference in remote observation detection between observation and non-observation trials. A mean effect size for the 11 studies was calculated resulting in a significant effect size (r = .25), and after careful consideration the authors determined that this considerable effect was not due to artefacts or confounding variables such as cueing, human error, deception, sensory leakage, or selective data reporting. The meta-analysis also showed that the mean effect size did not change whether the remote observer intended to raise or lower a participants' EDA through intention alone, or actual observation.

The second meta-analysis found a smaller effect size than that estimated by Schlitz and Braud (1997) and included all the same experiments; however Schmidt et al. (2004) reviewed additional papers - some of which had two-tailed analyses and thus counted both activation and calming as evidence of remote observation detection. The authors estimated that the collective mean effect size was weighted by sample size, and the result was a small, but significant estimated effect size (d = .28, the equivalent of r = .14). To take this is into account they applied a correction to the estimated mean effect size which resulted in a reduced effect size overall (d = .13, the equivalent of r = .05). Despite the differences in effect sizes, both Schlitz and Braud's (1997) and Schmidt et al.'s (2004) meta-analyses reported homogeneity of effect sizes across all of the studies they reviewed, and a significant overall effect was found suggesting a robust effect which was not due to outliers.

1.4.4 Does Scopaesthesia Increase Sympathetic Nervous System Arousal?

Braud et al. (1990; 1993a) conducted the first study using a psychophysiological measure and recruited 16 participants to be observed by his colleague Shafer, who was the observer in all trials and was separated by several rooms during the experiment to control for sensory leakage. Shafer placed electrodes on the participant, and explained that he or she would be watched during random periods using a CCTV camera, and would be visible via a monitor. She then went to the observation room, and opened a sealed list of randomised observation conditions to avoid accidental cueing during the preparation stage. Each participant had a sequence unique to them which was intended to rule out the possibility guessing patterns accidentally concurring with the order of conditions.

Even with such careful attention to detail, a significant increase in participants' sympathetic nervous system activation was found during remote observation trials. Importantly, this difference was during periods in which they were being remotely observed rather than an effect of the camera itself as this was running continuously during all trials. The effect size of $r = .57$ was larger than that found in the researchers' conscious guess studies, leading them to conclude that EDA is an effective measure of detecting remote observation.

Braud et al.'s (1990; 1993a,b) work was furthered by another study which found remote observation to be associated with increased EDA, however (Schlitz & LaBerge, 1994, 1997) also investigated whether the effect may be combined with the observer's intention effect the EDA of the observed participants. Schlitz and LaBerge's work incorporates findings from healing analogue studies in which remote influencers aimed to either calm or activate EDA responses through intention and imagery (Schlitz & Braud, 1997) as considerable support for the ability to influence a participant's physiology remotely was produced.

To see if this effect could be reproduced remotely, Schlitz and LaBerge had their participant pairs discuss ways in which they may be personally activated when under observation. They were then taken to different rooms which were 15 metres apart, with observation made possible by a CCTV link. To take into account that EDA may decrease over time, four blocks of trials were created: observed, not-observed, not-observed, observed, or vice versa. A computer programme ensured the sequence of trials was randomised and counterbalanced, and dictated whether the observer saw the participant or just a black screen. When the participant was visible to them, the observer attempted to 'activate' them and affect their EDA.

The 32 trials were 30-seconds long, and were interspersed with breaks of randomised duration throughout the experiment which controlled for sensory leakage through the use of separate rooms, and did not permit the experimenters to know the observation sequences until after any interactions with their participants in case of unintended cueing. Again, the observed participants' EDA was significantly higher during observation trials than non-observation, however it should be noted that the observers' deliberate intentions to activate their target may have caused the effect, rather than remote observation alone.

Further evidence that participants' EDA can be activated via remote observation was provided by Wiseman and Smith (1994a). Similarly, to their experiment with conscious guessing, also in 1994b, these researchers investigated whether a pair of participants separated by a screen could detect being observed by a group of individuals watching them via video cameras. It is this that differentiates Wiseman and Smith's studies from others which have used psychophysiological measures as their observed participants were not alone. During periods of observation, each of these participants were watched via CCTV by approximately half of the 13 observers recruited, who were located on an entirely different floor of the laboratory.

All participants rotated roles and acted as both the observed and the observer. Observers were instructed to either watch the images or look away by the experimenter who followed a pre-prepared randomised and counter-balanced sequence of observation trials. To create the sequence, a die had been rolled to establish a random entry into a random number table; the experimenters then assigned an observation trial proceeded by a non-observation trial to even, and vice versa for odd numbers. Each pair of participants were observed across six trials, with each trial lasting 20 seconds.

Results showed the difference between the EDA in the observation and non-observation conditions to be significantly different. However, the authors analysed the randomised and paired sequences and found that the pairing of observation trials before non-observation trials prevailed. They subsequently surmised that this could have created an artefact if the EDA of the participants being observed reduced over the course of the trials. Had there have been a genuine psi effect, this artefact may unfortunately have masked it. Had Wiseman and Smith used randomised ABBA counterbalanced blocks to balance practice effects, they would not have encountered this issue as the conditions would have been presented in one sequence, then the opposite.

To address the possible randomisation artefact, Wiseman et al. (1995) created another two follow-up studies to investigate EDA response to remote observation. This time the observed participants in both experiments took part individually rather than in pairs, and the rooms which separated the observed and the observer were in some instances approximately nine metres apart, and in others cases they whereon a different floor at the opposite end of the building. Observation trial randomisation varied in each of the experiments. The roll of a die determined entry into a random number table in the first experiment, with the following ten digits then used for the sequences. Even numbers translated to an observed trial followed by a non-observed trial, and vice versa for odd numbers. For every sequence obtained from the random number tables, a counterbalanced version was created listing the reverse observation conditions.

For the second experiment, the randomised ABBA sequences were employed for the ten observation and ten non-observation trials which lasted 30-seconds, and allowed participants a 30-second rest period between trials. Neither experiment produced a significant EDA activation effect, which could suggest that the positive results that Wiseman and Smith found in their prior (1994) research were indeed due to the possible artefact which they themselves identified, however their more recent non-significant findings could have been because of the experimenter effect.

An excellent example of the experimenter effect was found in a study conducted to investigate this well-known phenomenon, as well as the effect of remote observation on EDA. The experiment was created by Wiseman and Schlitz (1997) to examine the effects of a sceptic experimenter who had found non-significant results in their own remote observation research (Wiseman & Smith, 1994; Wiseman et al., 1995), with a believer in psi who had reported significant results from their remote observation experiments (Schlitz & LaBerge, 1997).

Randomised and counter-balanced observation sequences had already been prepared and it was only after the experimenter had left their participant that he or she learned of the observation sequence to avoid cuing or priming. Once the experiment began, the experimenter and the participants being observed were in different rooms separated by 20 metres. Observation was made possible by CCTV, and following a two-minute delay to allow the participant's EDA to settle to a baseline level, the observer either watched their image on a monitor, or looked away.

Trials were 30 seconds long and each participant underwent 32 trials. Even with stringent controls for sensory leakage and attention paid to avoiding cuing - when Schlitz was observing, the participants' EDA during remote observation trials was significantly higher than in the non-observation trials. Conversely when Wiseman was observing there was no significant difference between observation and non-observation in the participants' EDA.

These results certainly seem to suggest that psi proponents will elicit a psychophysiological reaction from those they observe, whilst a sceptic observer will not - however it is impossible to know for certain whether the difference in results was due to their interactions with the participants or their own psychic ability. The non-random allocation of participants to observers could have also resulted in a difference between the two groups which may account for the apparent effect of the experimenters observing - a risk which could have been avoided by comparing the two observers within-subjects. This study demonstrated however that remote observation can lead to an EDA response, and that the experimenter themselves may play a role in its detection.

The results of the EDA based experiments described above concur with spontaneous experience descriptions of remote observation, as they all found an increase in the sympathetic nervous system arousal of the observed participants which indicates the often-reported feelings of unease, nervousness and agitation. However, some studies have found remote observation creates a calming effect in those who are observed, and this will be discussed below.

1.4.5 Does Scopaesthesia Decrease Sympathetic Nervous System Arousal?

In a series of experiments which dictated that the observed and the observer participants worked together in connectedness exercises, Braud et al. (1993a, 1993b) found that remote observation had the effect of reducing sympathetic nervous system activity. In the first of these experiments (1993a) Braud and his colleagues used similar methodology, observation sequencing, procedure, and physical separation of the participants as in their experiment described in section 1.4.4 which led to evidence suggesting EDA can occur through remote staring. The exercises in connectedness however involved the participants sharing a single focus of attention while vocalising their present-time experience that generated a seemingly powerful feeling of interconnectedness (Andrews, 2004).

Prior to the observation trials, the participants who were to be observed joined together in the connectedness exercises along with Shafer, an experimenter who was also to play the role of the observer. Following the trials, the participants EDA levels were significantly lower whilst being observed than when they were not being observed. A possible reason for remote observation having a calming effect could be that the connectedness training had changed the "typical threatening" (Braud et al., 1993a, p. 387) feelings usually associated with scopaesthesia, and had been transformed due to the exercises building a more relaxed and comfortable relationship between the observer and observed. Interestingly, the connectedness exercises seemingly did not alter the magnitude of the remote observation effect, but merely changed the direction of its influence. It should be noted however that one should not necessarily assume that the connectedness exercises were the cause.

The same researchers conducted a further two experiments in which they also found that remote observation produced a calming response in participants (Braud et al., 1993b). The procedure for both experiments was the same as in their previous research on the topic (Braud et al., 1990; 1993a) except that in the first of these studies three new observers were trained by the experimenter Shafer, who had been the observer in previous versions. The new observers had the connectedness exercises explained to them, however they did not receive this training.

In the second experiment Shafer returned to his role of observer. Again, the results showed that the EDA of participants under observation was significantly lower in the observation trials than it was during the non-observation trials. The results remained even after checks had been made to ensure they could not be attributed to the participants being in coincidental synchrony with the randomised sequencing. A so-called 'sham control' was created in which the trials were still labelled as observation and non-observation sessions, however in reality they were all the same, with no observation taking place.

The trials showed no significant difference between the EDA of participants, and so offers a powerful counterclaim to the critique that previous evidence for remote observation detection is simply based on artefacts of naturally occurring rhythms which coincide with the order of observation sequences. It should be noted that the opposing observation effects before and after participants engaged in the connectedness training were not hypothesised by the authors. Braud et al. (1993b) proposed that the difference in direction from activation to a calming effect could be explained by a change in the attitude of the observer as during the first experiment, Shafer had been nervous and felt uneasy about his task - however she felt relaxed and comfortable during the second experiment.

It could be reasonably assumed that feeling more relaxed may have been a by-product of the connectedness training, but there is another possible explanation; that the observer felt more comfortable because of increased confidence gained from having already been part of the experiment. This familiarity with the remote observation process may have been transferred to the observers she subsequently trained, who could also have then exerted a calming influence. This notion suggests that the observer's emotions can be transmitted via observation, thus making it problematic to think of observation as a pure influence without contaminating factors. By conducting this series of studies, the researchers provided compelling evidence that monitoring participants' psychophysiological changes is an effective method of detecting remote observation, whilst highlighting that connectedness training may have the capacity to change the direction of remote observation influence from activating to calming.

The calming effect of remote observation was given further support following a study by Wiseman and Schlitz (1998) who replicated Wiseman and Schlitz's (1997) study into whether experimenter bias effects the study of remote observation. Using a similar methodology, the researchers opportunistically assigned their participants to two observers, who in turn followed the same procedure as the 1997 experiment. The observer first took the participants to be observed to the laboratory experimental room, which contained a CCTV camera, the EDA monitoring equipment, and headphones through which the participants would hear white noise to reduce the risk of noise cues. The observer then left to occupy a separate room a few meters away before starting a computer program created to randomise the sequence of conditions for the 30-second trials. The observer then watched the participants via a computer screen during the observation trials, but not during the non-observation trials.

As with their previous experiments, participants EDA only differed significantly between observation and non-observation studies in Schlitz's sessions. However, in their previous research Schlitz's participants showed a significant activation effect when under observation

(Wiseman & Schlitz, 1997), whereas in their 1998 version there was a calming effect. Unfortunately, the authors did not discuss this unexpected result or the possible reasons for it, however Braud et al. (1993b) suggested the emotions of the observer can be transmitted to those being observed.

It could therefore be that a similar effect occurred in Wiseman and Schlitz's (1998) study when Braud and his colleagues conducted follow-up research; namely that the observer was simply more relaxed during the second experiment. As the experiment was a replication, the procedures would have been familiar to Schlitz and so it is likely she would have been more comfortable for this reason alone, but when it is considered that the replication was of an experiment which had yielded a significant psi effect, Schlitz may have also had an increased confidence in her own psi abilities to produce an effect. Even carrying out the experiment at her place of work rather than at Wiseman's laboratory where the first experiment was conducted may have also contributed to feelings of familiarity, and therefore could help account for the calming effect of her observation. This possible explanation is of course highly speculative, but is potentially important to the design of such studies.

More recently in 2004, Lobach and Bierman contributed to the literature with a study which also focused on detecting remote observation, but this time the effect of informing or not informing the participants that they would be watched was considered by including the influence of a covert observer. In doing so, the researchers created the first experiment to measure remote observation detection when the participants were unaware that they would be watched - a condition referred to as non-intentional remote observation detection (Anderson, 2012).

To keep the remote observation elements undisclosed, participants in the Lobach and Bierman (2004) study were supposedly recruited for an experiment in which they would be asked to judge musical fragments. Their EDA and conscious guess measures were recorded whilst they rated whether 20 musical pieces of a 30-second duration were relaxing, by pushing a button on a response box. The participants were unaware that a hidden video camera was relaying his or her image to the observer via a computer screen in an adjacent room to avoid visual cues, although it should be noted that this did not deal with possible audio leakage.

The observer stayed within their room throughout the experiment, and did not meet the participants at any stage to negate any conventional experimenter effect. The sequence of observation and non-observation trials was randomised by the selection of paper slips with which each indicated a 'stare' or 'non-stare' condition, and a unique list was created for individual participants. Unfortunately, the order of observation conditions was not counter-balanced, which can be problematic in psychophysiology studies as EDA has a tendency to reduce over time.

Following completion of the first 20 trials, the experimenter entered the participant's room and explained that he or she would be remotely observed during the second half of the experiment at random intervals (Lobach & Beirman, 2004). A further 20 trials were then conducted lasting 30 seconds each, during which time the participant used a response box to indicate whether they believed they were being observed with their EDA measures being taken simultaneously. Whilst neither half of the experiment showed a significant difference between the participants' EDA during observation and non-observation trials, it is worth highlighting that an activating effect approached the significance level (p = .054) when the participants knew they were being remotely observed.

The experimenters interpreted these results as demonstrating that their participants' EDA responded to remote observation only when they were informed of its existence. However, with such a small and non-significant effect, this slight change could be due to random variation. There is also a possible confounding factor that must be taken into account, as during the first half of the experiment the participants listened to relaxing music - a likely result of which may have been the calming of their EDA responses. This could have over-ridden any activating effect of the remote observation, and unfortunately no investigation into whether the audio had influenced the participants' EDA was conducted.

1.4.6 Criticisms of Psychophysiological as a Measures for Detecting Scopaesthesia.

A major limitation of experiments which have used psychophysiological means to measure remote observation detection is that no other methodology has been used other than EDA. This method in remote observation studies has been criticised, with Schmidt and Walach (2000) claiming that the procedure used has not been up to current standards and so could either lead to artefacts, or cause experimenters to miss effects that could otherwise have been recorded.

This method of detection is also seemingly prone to the experimenter effect (Wiseman & Schlitz, 1997) and so results may be affected by the attitudes and beliefs of the researcher involved. There is also the added complication that even if there is a physiological reaction to remote observation, depending on the observers intentions and attention they could have the effect of reducing sympathetic nervous system activity, rather than increasing it (Braud et al., 1993a, 1993b). Whilst understanding this is a possible outcome should mean that it can be controlled for, there is no way an experimenter can truly know the intentions and attention of the observer. This suggests two-tailed hypotheses related to EDA are the most likely to capture significant effects, as sympathetic nervous system activity can be investigated for both increases and decreases.

1.4.7 Direct Comparison of Conscious Guessing and Psychophysiological Measures.

Despite neither the conscious guess measure nor the EDA measures resulting in a significant difference in their participants' responses to observation and non-observation, Lobach and Bierman's (2004) experiment provided some valuable insight into methods of measuring remote observation detection as it was the first of its kind to directly compare conscious guessing and EDA simultaneously. On average, the participants' EDA responses was one percent higher when they believed themselves to be under observation, as opposed to when they thought they were not being observed - this is in keeping with social facilitation which suggests that individuals are likely to have higher levels of arousal levels when they are proximately observed. This could therefore be considered as evidence that observation produces a similar effect, whether it is remote or proximate. It could however merely show that people are responding to their own arousal and *think* they are being monitored. An interesting direction for future research to take would therefore be to compare whether people react similarly when they *think* they are being observed, to when they know they are being observed.

Overall though, the experiments reviewed suggest there is substantial evidence that under the right conditions, remote observation can be detected. Furthermore, it seems that conscious guesses and unconscious psychophysiological measures both provide a means of detection which is significant and replicable. Behavioural responses may provide another method of detection though, as a change in behaviour may be influenced by both conscious and unconscious factors - and so task performance should theoretically be affected by remote observation. Whilst behaviour as an indicator of remote observation detection has been researched far less than either conscious guessing or EDA, spontaneous reports of turning to look at an observer (Braud et al., 1993a; Coover, 1913; Sheldrake, 2003; Titchener, 1898) and being woken from sleep (Sheldrake, 2003) would indicate that behavioural responses to remote observation are possible, and will therefore be investigated.

1.5 Studying Scopaesthesia via Behavioural Measures

1.5.1 Does Social Facilitation Have an Effect?

Social facilitation can be described as peoples' tendency to behave differently when alone, to when they are in the presence of others via either an increase, or decrease in their performance which can be attributed to being observed (Griffin & Kent, 1998). An increase in performance can be due to co-actions which can occur when an individual is either performing a task with others who are undertaking a similar task, or when they are performing a task in front of an audience. Everyday examples of co-action leading to social facilitation can be found when cyclist's perform at an enhanced level when part of a cycling team compared to when they cycle alone, or when a weightlifter competes in front of an audience and manages to lift more than they would do if the audience was not present. Such reactions can also be attributed to the Yerkes-Dodson law (Yerkes & Dodson, 1908).

The Yerkes-Dodson law is an empirical relationship between performance and pressure, which dictates that individuals' performance improves with mental and/or physiological stimulation - but only to a certain degree. When arousal levels become too elevated, the performance level decreases. When applied to social facilitation, this law states that the mere presence of other people will improve a persons' performance in speed and accuracy if a task is well rehearsed, but will decrease the performance of less practiced or more complex tasks. In other words, people tend to perform better on simple or well-practised tasks, and worse on complicated or unrehearsed tasks compared to when they are alone (Strauss, 2002).

A possible explanation for people performing worse on complicated tasks, and better if they find the task simple when in others' presence is because of their cardio-vascular response to the challenge. If performing a relatively simple task in front of another person or an audience, people tend to show a normal cardiovascular response, but if the task they are performing in the presence of others is difficult - their physiology responds as though they are under threat. Put simply, a normal cardiovascular response is likely to improve a person's performance, whilst a threat-like cardiovascular response is likely to degrade their performance (Strauss, 2002). With this in mind, it could be argued that the social element of scopaesthesia should at least be considered, as should the proximity and the familiarity of the observer.

Experiments investigating social facilitation and the Yerkes-Dodson law (Yerkes & Dodson, 1908) are methodologically problematic by their very nature. A literal 'alone' condition is essential to any related research to enable comparison between when a participant has an audience, and when they do not - however, a true alone condition requires a complete absence of presence, real or expected observation, and evaluation apprehension. In creating such a condition, an experimenter could not be present to ensure the experiment takes places as intended - indeed, even replacing the researchers with CCTV cameras to review the footage later to judge whether the study took place as planned means that a true alone condition is violated as the cameras themselves create an expectation of observation, and may even produce a similar observation effect to direct staring (see section 1.4.3). Alternatively put, a literal alone condition would successfully eliminate any influences that could create a social facilitation or reactivity effect.

Possibly due to the methodological difficulty with truly creating an alone condition, it has been either largely avoided, or poorly controlled for in the majority of research related to the topic (Guerin, 1993). This was demonstrated when Travis (1925) compared an observer condition of between four to eight students, and an alone condition with the just experimenter present. Whilst Travis did find a social facilitation effect from the observers, this does not mean that the presence of the experimenter exerted no social facilitation effect and therefore can only be interpreted that the observers generated a greater effect.

Similarly, Ekdahl (1929) found complex task inhibition from the observation of an experimenter in comparison to an alone condition meaning that experiments with poorly controlled alone conditions may fail to find social facilitation effect, or demonstrate it to lesser extent (Uziel, 2007). In either case, this could be due to an inadequate baseline, which could be used for comparison purposes. Unfortunately, these poorly controlled alone conditions have led to some discrepancies in the research literature as highlighted by Anderson (2012) who pointed out that Cottrell et al. (1968) declared that peoples' presence generated no greater effect than an alone condition - however the experimenter was listening to the participant in the alone condition, which creates an evaluative influence. As mentioned earlier, this is problematic for a true alone condition.

Through improved alone conditions, experimenters (Markus, 1978; Schmitt et al., 1986) have managed to demonstrate a social facilitation effect from mere presence which is a possible experimental issue for scopaesthesia research, but the researchers concluded that influence is likely to be weak when compared to proximate observation. It could be argued however that this problem is virtually impossible to circumvent as the issues with creating a true alone condition even when the tasks under observation are covert in nature are difficult to overcome as the participants are likely to 'feel' monitored, regardless of if they are genuinely under observation or not (Griffin, 2001; Griffin & Kent, 1998).

Griffin and Kent (1998) made the prediction that participants performing a task would assume that monitoring would be taking place, and that even in the absence of any evidence that this was the case that they would imagine the means by which they were being observed. To test this, they allocated participants randomly to one of three conditions: alone with no task; alone with a task involving card sorting; and alone with a different card sorting task which was self-timed with a stopwatch. The researchers found that the participants working on a task felt significantly more concerned about their performance, and were more convinced that they were being judged and monitored than the participants who were not performing a task.

Griffin (2001) used a follow-up study to compare the same three alone conditions as Griffin and Kent (1998), along with a proximate observation condition. Interestingly, the participants felt the most monitored under proximate observation, followed by those completing the tasks. Those in the alone condition, but without a task felt the least monitored. This logically leads to the conclusion that participants engaged in undertaking a task will imagine that they are under observation regardless of whether this is the case - even without any apparent means of the experimenters observing them, and that they will subsequently experience the associated evaluation apprehension, despite the fact that they were alone. It is therefore reasonable to conclude that physical aloneness does not necessarily equate to phenomenological aloneness, and that this research should be considered when conducting studies involving tasks or problem solving. This is particularly relevant if the task is not covert, and the participant is aware of it taking place.

An example of an experiment which used a covert remote observer is Markus's (1978) work in which each participant was asked to change clothes under one of three observation conditions, to which they had been randomly assigned. In these conditions they were either i) watched by a confederate situated within the same room (the proximate observation condition), ii) joined in the room by someone with their back turned to the participant as they pretended to be busy fixing machinery (the mere presence condition), or iii) physically alone in the room (the alone condition). In every condition the participants were timed by a researcher watching through a one-way mirror. Markus found a social facilitation effect was found, with a significant increase in the time the participants took to change into unfamiliar clothing from 'alone' to 'mere presence', and 'mere presence' to 'proximately observed conditions'. There was also a significant decrease when this was compared to the participants changing into familiar clothing. This indicates that the 'alone' condition was indeed a sufficient baseline to compare with the 'mere presence' condition. It should be noted however that Markus's decision to use a covert, remote observer in the alone condition indicates that she did not predict that such a form of observation would be influential.

A social facilitation experiment which did use a true alone condition in conjunction with covert tasks took place when Schmitt et al. (1986) attempted make improvements to Markus's (1978) experiment. Schmitt and his colleagues criticised Markus's idea of having her participants change into ill-fitting clothes with the intention of causing embarrassment and anxiety about their appearance. Instead, the researchers employed typing tasks, which they did not predict would cause embarrassment or apprehension. For the simple task, participants were instructed to type their name before undertaking the complex task, which was typing their name backwards, interspersed with ascending digits. Both tasks were presented as taking place prior to the experiment beginning, which was introduced as being about sensory deprivation. The participants were assigned randomly to one of three observation conditions: namely mere presence, proximate observation, or alone.

In the 'mere presence' condition, a blindfolded confederate wearing head-phones and therefore unable to see or hear was situated in the room with the participant. In the 'proximate observation' condition the experimenter was stood in the same room as the participant, noticeably looking over their shoulder. Finally in the 'alone' condition the participant was alone in the room and unobserved. None of the participants were aware that their typing speed was monitored and recorded by the computer, and therefore negating the need for any human observation to measure performance or conduct. The experimenters found that none of the participants suspected any monitoring was taking place, or that the tasks formed part of the experiment. A significant and critical pattern for social facilitation was found, as when compared to the alone condition both of the 'audience' conditions were completed quicker for the simple task, and slower when undertaking the more difficult task. The researches therefore claimed this as evidence that mere presence is adequate to produce social facilitation.

The tasks employed in Schmitt et al.'s (1986) experiment did not cause participants to assume that they were being monitored, observed or evaluated as they were presented simply as preparation for the experiment itself. The time the participant took to type was recorded without the need for human intervention which is ideal for a social facilitation experiment that uses a remote observation condition. Furthermore, deceiving the participant into believing that they were not being observed may have played an important role in reducing anxiousness. In this sense the experimenters were successful in creating an alone condition which allows comparison to remote observation conditions.

With the mere expectation of observation shown to affect individuals' behaviour, social facilitation should be considered within scopaesthesia research both as a possible indicator of remote observation detection *and* as a possible barrier to effective experimentation which must either be overcome, or at least accounted for. To investigate this potentially important and interesting measurement of remote surveillance detection further, it was included as part of the current research.

1.5.2 Behavioural Measures as a Means of Detecting Scopaesthesia.

The most often used form of detection in remote observation studies appears to be conscious self-reports. This is likely to be due to the ease with which they can be set up, conducted, and analysed. For this reason, such a measure should ideally be utilised in any research on the subject if the experimenter intends to be able to compare and contrast the study with those contained within the literature - however behavioural measures should also be fully explored as this is under-researched when compared to either conscious guessing, or psychophysiological response.

Sheldrake's (2003) experiment was designed so that six observers were hidden from the observed participants by a one-way mirror. The observers were located in an office overlooking the BBC studio entrance in which a sizable audience was waiting to enter. All audience members had given permission for video images of them to be recorded, but they had in fact been filmed already. Sheldrake's assumption was that during the period of actual observation and covert filming, the audience members were not expecting this to be taking place.

Together, the observers all stared at the audience members at the same time, or all looked away in unison. These conditions were pre-prepared according to a randomised schedule. All of the participants being observed were facing away from the one-way mirror during these crucial moments. The observers' office housed a video camera that continuously recorded the audience waiting outside, and the number of times the audience members turned to face the camera was later counted and recorded by a judge who was unaware of the order or duration of the observation and non-observation periods. It should be noted that turning to look at the observer is among the most often reported behavioural responses to being stared at.

Overall, the participants turned to face the camera 27 times during the observation periods, and just 12 times during the non-observation periods. This is a highly significant difference in the reactions of the people under observation. Should such a difference represent a genuine psi effect, the influence exerted by the hidden observers direct staring was greater than the effect of the video cameras' constant electronic observation. Interestingly, when the audience members were interviewed about the experiment, none stated an awareness of having turned around, and none reported feeling as though they were under observation. This absence of awareness suggests that those under observation responded automatically, rather than via a deliberate action - a finding that hints that scopaesthesia can influence behaviour, and that direct staring will elicit a more salient response than electric surveillance.

Sheldrake's (2003) study tested behavioural response to hidden observers, but Lee at al. (2002) had already considered behaviour as a measure of detecting remote observation. During this earlier study, the observed participants were informed that they would be randomly observed whilst performing a coordination task which involved them passing a metal loop along an awkwardly shaped wire. Should the participant make a mistake and touch the loop against the wire, an electric circuit was closed, and a warning tone would sound; an error would subsequently be recorded. The process was observed in an adjacent room via CCTV, although the position of the camera was varied between being in front or behind the participant. Additionally, the number of observers varied also between one and three people. However, upon completion of the experiment - observation, the number of observers, nor the camera position appeared to have made a significant difference to the coordination task error rate. Unfortunately, Lee and his colleagues did not consider the social facilitation effect that non-remote observation may have on task performance, and so could not predict whether remote observation would affect behaviour, or whether performance would improve or deteriorate.

Whilst both Sheldrake's (2003) study and Lee et al.'s (2002) research attempted to incorporate behavioural responses into the body of work related to remote observation, future experiments would benefit from considering the social facilitation effect - in other words the effect that normal observation has on a person's performance. Ideally, this method of testing for detection would be combined with the more established conscious guessing and psychophysiological measures so that the results can be compared and contrasted to offer a unique approach that stands the best possible chance of capturing the phenomenon of scopaesthesia, should it indeed exist.

1.5.3 Could Scopaesthesia elicit a Social Facilitation Effect?

If people can respond via extrasensory means as they would do in the same circumstances given sensory knowledge of the situation (Stanford, 1990), it should therefore follow that if normal sensory awareness of observation (proximate observation) leads to the social facilitation effect, remote observation could elicit the same reactions. Behavioural responses should therefore form at least part of any modern research intended to further the field. Stanford (1990) also hypothesised that changes in a person's sympathetic nervous system would occur in response to specific circumstances, so perhaps this may be true of a physiological change from remote observation, if a similar change occurred from proximate observation.

It should also be considered that remote observation can influence the sympathetic nervous system of the person being observed (Braud et al., 1993). The social facilitation effect can be mediated by changes in arousal, with activation being the most likely effect, unless there is a specific reason for a calming effect. Whilst this is still not accepted universally, it remains the primary explanation of the social facilitation effect according to Mullen et al. (1997). It could therefore follow that if remote observation is associated with changes in the sympathetic nervous system, and these changes lead to the social facilitation effect - theoretically remote observation would also lead to the social facilitation effect. Additionally, remote observation might elicit a social facilitation effect if the person being remotely observed becomes aware of this as such circumstances can influence behaviour in the same way as when the observation is proximate (Sheldrake, 2003).

Although psychology-based investigations have been inspired by a variety of spontaneous experiences reported over the last century in an attempt to establish whether remote observation detection may be a genuine phenomenon, doubt remains as to its validity. The initial studies in this controversial field (Coover, 1913; Titchener, 1898) were based on poorly controlled experimental designs using conscious guessing as the measurement of detection, which have developed over time to include more sophisticated and carefully managed versions. Many of these have suggested an effect from remote observation detection, and studies utilising unconscious psychophysiological measures appear to capture this effect even more reliably.

It could therefore be surmised that behavioural responses may be an effective method of detection when investigating remote observation as behaviour can respond to both conscious and unconscious influences, yet this remains under-researched. One possible reason may be due to the established theory of the social facilitation effect (the effect of normal observation on task performance) and how relating it to remote observation could incur difficult methodological issues - this will be considered in detail throughout the following section.

1.5.4 Methodological Considerations When Investigating Social Facilitation Effect.

It is particularly important in psi related research that sensory leakage, inadvertent cueing, predicable condition ordering, and non-random participant allocation is avoided as this weakens the data gained from the study and can invalidate any conclusions drawn from the results. In addition, there are further aspects to consider when aiming to research the interaction between an observer and an observed participant with respect to the social facilitation effect.

One of these methodological issues concerns the behaviour of the observer whilst proximate observation is taking place, as research that compared continuous staring with intermittent observation found the former elicited a greater social facilitation effect (Huguet et al., 1999). Similarly, the observers' location has been shown to make a difference as observers who stand diagonally behind the individual being observed so as to look over their shoulder have exerted a greater social facilitation effect than observer's who are positioned in front of the participant

(Huguet et al., 1999; Klauer et al., 2008). Experimenters should therefore consider this potentially important experimental aspect when designing their study.

Additionally, familiarity between those involved should be taken into account as limited findings suggest this can affect the observed participants' behaviour differently (Guerin, 1993). According to Bond and Titus (1983), if the pair of individuals are familiar with one another, the social facilitation effect will be reduced. Familiarity's role in social facilitation would appear to make sense when considering the sympathetic nervous system. Whilst this explanation is not unequivocal, it can be interpreted that remote observation increases an individual's sympathetic nervous system activation, but that connectedness can reduce such activation as demonstrated by (Braud et al., 1993a, 1993b). This is a tenuous link; however, it suggests that it would be prudent for researchers to recruit participants who are strangers, avoid deliberately or accidentally creating bonds between them, or that this should at least be considered and controlled for.

1.5.5 Should Behavioural Measures Be Considered?

The literature demonstrates and highlights empirical findings from previous studies relevant to both scopaesthesia and the social facilitation effect, and indicates that various forms of social presence can influence this effect; namely proximity and familiarity. It could be argued that conventional observation can produce the social facilitation effect, and that everything learned thus far from the study of this phenomena could and should be applied to research relating to remote observation detection to further examine the possibilities of its very existence, as well as predictive factors. Real-world applications related to whether covert observation makes any difference to an individual's behaviour includes, but are not limited to whether terrorists or criminals may cease the activity for which they are under surveillance when being covertly watched, whether managers will see their employees typical behaviour when observing them, and whether testing individuals under observation can truly be reflective of their everyday ability.

Whilst scopaesthesia research has shown substantial evidence that such abilities may exist, despite the notable experiments of Lee et al. (2002) and Sheldrake (2003), behavioural measures as a means of detection are under-researched and so meaningful conclusions cannot yet be drawn. Incorporating behaviour and performance into the body of research by considering the well-researched social facilitation effect could provide a valuable extension to both fields of study.

A review of the literature and past findings in relation to observation effects from the fields of remote observation detection and social facilitation show that various forms of social presence have been shown to lead to the social facilitation effect. It could be argued that direct observation elicits the social facilitation effect, and therefore testing for the social facilitation effect via remote observation would further examine the claim, as well as exploring whether there is a justification for covert observation if it makes no significant difference to task performance. Whilst evidence has been found to support the notion that unconventional detection of remote observation may be possible, behavioural measures are currently in need of further research as the existing literature is inconclusive.

Considering whether or not remote observation could cause behavioural change would improve and extend both fields of research, and by combining the heavily-researched area of social facilitation from proximate observation and the elusive influence of psi, answers to previously unanswered questions may be found. In turn, this may inform the design of not only experiments related to scopaesthesia, but indeed all studies that requires observation of any kind.

1.6 Factors to Consider When Studying Scopaesthesia

1.6.1 Individual Differences in the Detection of Scopaesthesia.

As robust evidence would provide support for the validity of extrasensory ability; the research will also investigate psychosocial and neurological predictors which may be related to the phenomena to assist with future studies and more targeted sampling. With Williams (1983) experiment considering whether participants' belief in psi may play a role in their ability to detect remote observation, Lobach and Bierman's (2004) finding that the observer's belief in psi is a relevant variable, and Wiseman & Schlitz's (1997, 1999) research reinforcing the idea that results can be affected by the belief of the experimenter - evidence suggests that anomalous belief should be included in remote observation experiments as a predictor of detection.

To further explore whether psychosocial and neurological predictor variables are associated with the ability to detect covert surveillance, the current study considered additional variables; namely temporal lobe lability and schizotypy. The rationale for the inclusion of these factors is discussed below.

1.6.2 Experience and Belief as a Predictor of Surveillance Detection.

Belief in psi is shared by the majority of the population (Ross & Joshi, 1992) with paranormal experiences being reported worldwide (Sheils, 1978). Whilst this in no way means that such occurrences are a genuine phenomenon, it does indicate that the topic is worthy of investigation. Research suggests the paradox that such beliefs can be associated with improved mental health (Shumaker, 2001), whilst admitting to them can lead to negative stereotyping (Holt et al., 2008). Whether a positive or negative trait, belief in psi has been linked to paranormal experience generally (Wiseman et al., 2003), and to remote observation detection specifically (Lobach & Bierman, 2004; Williams, 1983; Wiseman & Schlitz, 1997, 1999).

Lindeman et al. (2008) attempted to explain belief in psi as a misunderstood domain specific knowledge which is related to physical, mental, and biological occurrences, and subsequent attempts have been made by researchers to create subscales of paranormal belief (Tobacyk &

Milford, 1983), as well as various types (Irwin, 1997) to investigate it further. Belief in psi as a measurable individual difference related to psi ability is potentially important as it has been proposed that such belief can determine whether an otherwise ‘normal’ event is interpreted by the individual as a paranormal experience, or simply no more than a coincidence (Bressan et al., 2008; Broughton, 1988; Falk, 1989; Metzinger, 2005; Stanford, 1974).

Whilst belief in psi is on a continuum, experimenters often try to establish dichotomous groups - non-believers who are pre-disposed to rational thinking and are likely to consider factors such as probability or environmental factors, and believers who are more likely to attach meaning to their experience (Alvarez, 1965; Bering, 2006). Individuals in either group may experience what they expect to happen, based largely on their pre-existing beliefs and experiences. Those susceptible to letting expectation determine their interpretation are not necessarily deliberately and consciously doing so, as self-deception strategies may be playing a role (Hergovich et al., 2010), and so it could be unintentional (MacCoun,1998) - however these cognitive biases (Kahneman & Tversky, 1972) have been demonstrated to share a relationship with psi ability (Kennedy & Kanthamani, 1995; Kohr, 1980; McClenon, 1994; Palmer, 1979; Ring, 1984).

Respected researchers such as Lange and Houran (1997) provided further support for the idea that belief in psi and psi ability are linked when they conducted their experiment into this possible relationship during which participants were told a location was, or was not haunted, with a primary focus on internal perception and the sensations their experience evoked. The results led the authors to conclude that expectation and belief are the most vital single factors which contribute to a paranormal ‘experience’. The researchers found further evidence for this in subsequent studies (Lange & Houran, 1998; Lange & Houran, 1999; Lange et al., 2000), however, in Laythe and Owen’s (2012) study, neither the Anomalous Experiences Inventory (Gallagher et al., 1994), or the Paranormal Belief Scale (Tobacyk, 2004) significantly predicted haunting experiences.

Whilst individual experiments have found evidence for and against the relationship between the belief in psi, and psi abilities - a considerable amount of data exists which supports this link. One of the arguments against the persuasiveness of this research relates to the lack of replicability when it comes to significant findings. In an effort to address some of these concerns and to gain a better understanding of which factors could be associated with more powerful and consistent effect sizes, Zdrenka and Wilson (2017) analysed all previous experiments which had focused on forced-choice precognition and had analysed individual differences. The researchers aggregated measures such as personality traits to determine which factors may predict psi performance most reliably. A total of 55 studies published between 1945 and 2016 were included, and this meant that no less than 35 individual difference measures were subject to meta-analysis - 6 of which (the belief that luck is controllable, openness to experience, perceptual defensiveness, time belief as dynamic, extraversion, and belief in psi) were found to be significantly correlated with psi performance. Indeed statistician Utts (1991) claimed that a "promising direction for future process-oriented research is to examine the causes of individual differences in psychic functioning" (p. 377).

Zdrenka and Wilson's (2017) meta-analyses suggest that such an approach is necessary, given individual difference factors such as belief in psi have been thoroughly investigated and thus suggest a promising avenue of research with regards to this topic - especially as many researchers have ignored individual difference factors, and in doing so may have missed potentially important sources of between-individual variation in performance related to psi. The authors themselves suggest that an actual effect could possibly be hidden if an individual difference factor has a systematic relationship with psi performance. For instance, if high-scoring participants on a certain trait over-perform, whilst low-performing participants in the same trait underperform - they would effectively nullify any effect and cancel each other out. For this reason, the meta-analysis was conducted to synthesise the existing research in an effort to understand which factors, if any, may lead to a convincing demonstration of psi in laboratory-based environments.

There had previously only been one meta-analysis conducted looking at belief in psi, also known as the sheep-goat effect (Lawrence, 1993) which found a relationship with psi performance (r = .03) suggesting a modest, yet robustly significant overall effect size. Zdrenka and Wilson's (2017) own meta-analysis demonstrated that belief in psi was the most often studied individual difference correlate in experiments utilising forced-choice responses, with this individual difference having been reported on in 22 studies by 12 independent researchers investigators who had based their results on the responses of 2,200 participants overall. The authors found correlations which ranged from -.17 to .72, and an overall mean weighted effect size (r) of .13 (p = .002). This indicates a significant relationship between performance on a psi task and the participants' belief in psi - and subsequently suggests that individuals who believe in psi are more likely to perform better when undertaking activities that relate to psi when compared to those who do not believe in its existence. This effect size is marginally larger than the effect size (r = .03) Lawrence (1993) reported in his meta-analysis based on the sheep-goat effect - however it should be noted that this also included clairvoyance and telepathy experiments. Researchers basing their own experiments on these findings would be wise to consider that this latest meta-analysis also found that a test of heterogeneity was significant ($p < .001$) suggesting that the variation in results could be explained by factors other than those under investigation.

These meta-analyses are important as without them, experimenters may impose their own individual synthesis of their data, and they are also able to provide greater insight even when they only incorporate a few studies. Indeed, Valentine et al. (2010) have argued that no other technique provides such transparency, and that all other alternatives are likely to be less valid. Zdrenka and Wilson’s (2017) meta-analysis results are also consistent with Steinkamp’s (2005) review of such experiments which led her to conclude that “there are few variables which have correlated clearly with success... most variables tested provided little evidence either way as being ultimately psi-conducive and there were relatively few variables that appeared to be encouraging” (p. 155). There are however, two notable exceptions according the meta-analysis results - extraversion and belief in psi. These individual differences demonstrate more consistent results, and do so across an impressive number of studies.

In their paper, Zdrenka and Wilson (2017) attempted to explain the apparent relationship between extraversion and belief, and their apparent link with psi performance by suggesting that there could perhaps be a mechanism derived from open-mindedness, curiosity, and intuition. The authors theorised that these personality traits may lead those who score highly in them to consider, discuss, and explore unusual ideas - they also extrapolated from this idea that these same individuals may consequently act on intuitions or information when others may suppress or ignore them. The notion that creativity and imagination may be correlated with psi belief and ability has also been explored by Groth-Marnat and Pegden (1998), Irwin and Green (1999), Kennedy (2005), Lindeman and Aarino (2006), and Smith et al. (2009) who all found a positive relationship in their data - thus adding to the literature which supports the association between belief in psi, and psi ability.

According to Palmer (1979) and Blackmore (1986), there is a positive correlation between belief in psi and psi experiences, with similar or equivalent research methodology employed in investigating them. They therefore suggest that belief in psi may be a result of perceptually experiencing it, and vice versa. Whilst the current experimenter acknowledges that previous studies have found opposing results, there is sufficient evidence that belief in, and experience of psi may be a predictor of psi performance - participants taking part in the current studies were therefore asked questions relating to their level of anomalous experience and belief to investigate whether this is related to their covert surveillance detection ability.

1.6.3 Temporal Lobe Lability as a Predictor of Surveillance Detection.

Parapsychology research focusing on the temporal lobe dates back over a century as the idea that this could be related to paranormal experience is not new (Spratling, 1904). Indeed, patients displaying partial seizures with foci within this area of the brain have reported Unusual Experiences for several decades (Gloor et al., 1982). It has been argued that such temporal lobe symptomatology could be responsible for a variation of apparent anomalous experiences (Persinger, 1984), and that temporal lobe lability may explain some mystical and religious experiences (1983, 1984).

Further evidence that there could be a relationship between the temporal lobe and supposedly psi related experiences can be found via direct electronic stimulation of this region to create sensations usually associated with encounters assumed by the experiencer to be paranormal (Horowitz & Adams, 1970; Persinger et al., 2000). Following extensive research, it has been suggested a temporal lobe sensitivity continuum exists along which all individuals are distributed (Cook & Persinger, 1997), and that people who give comparatively high (above-average) responses to Persinger and Makarec’s (1993) Personal Philosophy Inventory are likely to report more frequent paranormal experiences as this measure was designed to evaluate signs of temporal lobe lability.

The lability of an individual’s temporal lobe could begin to explain why some people interpret and decipher identical information in very different ways, however it could be especially important to the current study as the superior temporal gyrus incorporates an area where external auditory signals first reach the cerebral cortex region - this area is the primary auditory cortex and is engaged during the process of listening (Squire et al., 2004). The adjacent areas are known as the superior, posterior and lateral temporal lobe regions and they are essential for high-level auditory processes. With unexplained sounds featuring heavily within the reports of anomalous experiences, researchers have concentrated their focus on this fascinating area of the brain. Subsequently questionnaires and inventories have been created to measure temporal lobe lability, thereby negating the need for MRI scans and allowing researchers to avoid the costs and risks associated with them in the process (Neppe, 1983). Such questionnaires and inventories have been utilised in research which has demonstrated a relationship between temporal lobe lability and paranormal belief and experience (Irwin, 2009; Luke et al., 2013; Wiseman & Greening, 2005).

Dewhurst and Beard (1970) extended their research investigating the link between anomalous experiences and epilepsy when they analysed brain scans from patients who were suffering from seizures throughout their lives. As well as other emotional distresses, all of the six patients reported significant anomalous experiences at some point during their seizures which were religious in nature - and most detailed a sudden and unusual feeling of clarity and elation. X-rays and air encephalograms showed evidence of temporal lesions in four of the six patients, and electroencephalograms (EEGs) revealed spiked discharges of electrical activity in the temporal lobe in all six patients. This increased temporal lobe activity combined with the deformities found in two thirds of the patients provided a significant correlation between anomalous religious experiences and the temporal lobe.

During their discussion of the results, Dewhurst and Beard (1970) considered varying approaches to the issue of such anomalous experiences in epileptic patients. Their explanation referenced Jackson (1876), who posited that the cause could lie within temporal lobe abnormalities such as the electrical discharge which may simultaneously cause a loss of functioning in the brains' higher regions (such as the cortices) and hyperactivity in the brain's lower regions (such as the limbic system). The researchers argued that such a two-sided effect could diminish the important cognitive process of reasoning, however at the same they suggested that the lower brain functions would be time agitated. Whilst this seems to suggest that religion, and by association, other anomalous experiences are a lower psychological brain function - this is perhaps better described as the brains' inability to interpret the information carried from its lower regions. Although Jackson could not employ the technology to test his theory empirically, he suggested that such confusion within the brain creates an ideal psychological environment for rationalizing perceptions fostered by past experience.

Ramachandran (1998) offered a further explanation of such anomalous experiences and detailed how repeatedly electronically stimulating the brain can facilitate pathways between neurons, or even open new ones. This process is called ‘kindling’, and it was developed to explain how repeatedly stimulating a specific brain area can elicit the onset of seizures, and that in some cases this could continue even after the controlled stimulation had ended. Goddard et al. (1969) found that the threshold for activation of a seizure lowered following repeated electrical stimulations to specific areas of the brain.

According to Britton and Bootzin (2004), many studies suggest that an altered functioning of the temporal lobe, particularly within the right hemisphere is responsible for - or at least related to anomalous and religious experiences. These researchers investigated temporal lobe functioning in participants who had reported experiencing transcendental near-death experiences resulting from life-threatening situations. The individuals studied showed more temporal lobe epileptiform electroencephalographic activity than those within the control group to which they were compared. They also reported significantly more temporal lobe epileptic symptoms.

Contrary to the results predicted by the experimenters, epileptiform activity was almost completely lateralized to the left hemisphere. The study showed that near-death experience did not appear to be associated with dysfunctional stress reactions such as posttraumatic stress disorder, dissociation, and substance abuse - but rather was related to positive coping styles. Further analyses demonstrated that near-death experiencers typically slept for a shorter duration and had delayed REM (rapid eye movement) sleep when compared to the control group. This data suggests that altered temporal lobe functioning could be involved in the near-death experience, and hints that those individuals who have experienced the phenomenon are distinct physiologically from the general population.

Due to the long-standing association between temporal lobe lability and paranormal encounters (Horowitz & Adams, 1970; Luke et al., 2013; Persinger et al., 2000; Roney-Dougal et al., 2014), lability of the temporal lobe was considered as a possible predictor of surveillance detection in the current study. For details on how this will be measured and its accuracy, see section 2.2.3. Whilst the neuro-anatomy of emotional and cognitive interactions in individuals experiencing temporal lobe seizures was proposed over a century ago (Clouston, 1892), it has since been suggested that neuropsychologists should perhaps focus on uncovering issues concerning the extent to which temporal-limbic hyper-connectivity may account for schizophrenia-like epilepsy

(Weber et al., 2006; Pessoa, 2008); in doing so they may find empirical links between temporal lobe lability and schizophrenia - a mental disorder usually characterized by poor emotional responsiveness and a breakdown of cognitive processes (Kapur, 2009). A precursor of this condition is believed to be high levels of schizotypy, and this possible predictor will be discussed in section 1.6.4.

1.6.4 Schizotypyas a Predictor of Surveillance Detection.

As a multi-factorial personality construct, schizotypy shares a substantial variance with temporal lobe lability (Simmonds-Moore, 2010), and exists on a continuum with psychosis (Claridge, 1997; Eckbald & Chapman, 1983). Whilst the researcher conducting the current study is open-minded to the apparent effects of unusual phenomena, such beliefs are often associated with negative traits such as a suggestibility, an inability to think critically, to mental health, or even psychopathology according to the literature (Holt et al., 2008). However, belief in the paranormal has been found to be highly correlated with schizotypy (Simmonds-Moore, 2010); especially that of the 'healthy, 'positive' or 'happy' schizotype' (McCreery & Claridge, 1995). Whilst individuals belonging to this schizotypal sub-group are particularly prone to reporting Unusual Experiences such as pseudohallucinations, it has been shown that they may exhibit no indication of a mental disorder, even though they report above average schizotypy scale scores.

Typically experiencers of the paranormal, or those who believe in its existence belong to this so-called 'healthy' schizotype category (Mason et al., 2005), which may be explained by the unique way in which these Unusual Experiences are evaluated, as it could be that unhindered by the Cognitive Disorganisation associated with individuals scoring *exceptionally* highly on the schizotypy scale (Schofield & Claridge, 2007), 'healthy' schizotypes may be more likely to perceive their Unusual Experiences from a positive perspective. Such positivity potentially leads to imaginative tendencies (Lynn et al., 1996) and creativity (Bak et al., 2003; Brod, 1997; Claridge & Beech, 1995; Holt et al., 2008; Jackson, 1997; Nettle, 2006). Individual papers indicate that schizotypy has a relationship with both beneficial and detrimental traits, and cluster analyses have also suggested two types of high positive schizotypy scorers with one relating to positive mental health, and the other experiencing negative effects. This challenges the idea that there is a clear and direct association between unusual beliefs and experiences and psychopathology.

Individuals who suffer schizophrenia tend to exhibit higher scores on the psi and superstitious subscales of the Revised Paranormal Belief Scale (Shiah et al., 2014), compared to healthy individuals. Some researchers argue that different facets of schizotypy relate to different components of psi beliefs (Hergovich et al., 2008; Hergovich et al., 2005), whereby the cognitive-perceptual component of schizotypy exhibits a stronger association with paranormal belief than the interpersonal or disorganised components of schizotypy. Mathijsen (2016) argues that schizotypy is strongly associated with anomalous experiences, but suggests that the relationship between schizotypy and psi belief is heavily mediated by cognitive processes and change of world views (i.e. paradigm shifts). In addition, Barnes and Gibson (2013) argue that supernatural or paranormal experiences are strongly correlated with higher scores in positive schizotypy, which further supports evidence to suggest that different facets of schizotypy are correlated with different aspects of psi belief (Hergovich et al., 2008).

Schizotypy as a term is derived from 'schizophrenic genotype' and suggests an enhanced disposition toward schizophrenia (Claridge, 1997). According to the research literature, it is currently understood in two forms - advocates of a quasi-dimensional approach such as Eckblad and Chapman (1983) propose that schizotypal traits exist on a dimension, but suggest their presence is indicative of possible future psychopathology. Other researchers such as McCreery and Claridge (1995) endorse the fully dimensional approach and interpret schizotypy as a personality continuum upon which all people vary. In this second model, schizotypy interacts with risks such as stressful events and protective variables such as supportive social networks - which in turn leads to positive or negative outcomes, such as creativity or psychosis (Bak et al., 2003; Brod, 1997; Claridge & Beech, 1995; Jackson, 1997; Nettle, 2006) and so is neutral in terms of mental health.

Relatively recent studies have established direct support for the fully dimensional approach, rather than the quasi-dimensional approach. This suggests that schizotypy is more likely to be a personality dimension rather than a pre-cursor to diminishing mental health (Rawlings et al., 2008). However, this supporting evidence is derived largely from studies that found a relationship between schizotypy or Unusual Experiences and well-being (Goulding, 2004, 2005; Kennedy et al., 1994; Kennedy & Kanthamani, 1995; McCreery & Claridge, 2002). Such results are supported by the high proportion of individuals from non-clinical populations who report Unusual Experiences as approximately 10% of the general public have had one or more non-drug-induced hallucinatory experiences according to surveys (Bentall & Slade, 1985; Posey &

Losch, 1983; Sidgwick, 1894). Such hallucinations are more frequently experienced by individuals in stressful or traumatic situations such as bereavement or having been deprived of sleep or their senses (Bentall, 1990; Lukoff, 2007; West, 1962).

Further supporting evidence for the fully dimensional approach can be found in the variation of intensity of hallucinations as more commonplace perceptual experiences are considered 'normal' such as hypnagogia and daydreaming (Bentall, 1990; Fosse, Stickgold & Hobson, 2004). Similarly, subjective reports of parapsychological experiences are not rare across the general public, worldwide. For instance, telepathy (the ability to communicate directly between one mind and another) has been experienced by 33% to 50% of survey respondents cross-culturally (Glicksohn, 1990; Palmer, 1979; Rice, 2004; Targ et al., 2000).

Gianotti et al. (2001) propose that such anomalous occurrences may be explained by cognitive disinhibition and suggest a continuum of associative processing ranging from creative thinking, through paranormal ideation in healthy individuals to disordered thought processes, psychopathological delusion, or apophenia (the "specific experience of abnormal meaningfulness" (p. 596). It has also been considered that Unusual Experiences exist on a continuum (Bentall, 2000), and that similar biases in the way individuals' process information may contribute to unusual beliefs, pathological hallucinations and delusions.

Regardless of the approach, schizotypy is agreed to be a multi-factorial construct (Mason et al, 1997) comprising of four core traits. The most consistent factor within these traits is 'Unusual Experiences' which includes religious or magical beliefs, altered perceptions and sensations and perceptions of an individuals' own body and the world around them, hypersensitivity to the eternal environment including smells and sounds, a sense of déjà vu, jamais vu, as well as auditory and pseudo-hallucinations (Mason et al., 1997). The symptoms associated with this first factor are thought to be typical, and underpinned by disinhibition at the cognitive and/or sensory level, or weak gating (Claridge & Davis, 2003) which it is suggested could lead to flooding of the individuals' contents of consciousness (Burch et al., 2004; Frith, 1979; Gray et al., 2002).

Mason et al., (1995) suggest that a second factor, namely 'Cognitive Disorganisation' which can reflect difficulties with decision making, attention and concentration in conjunction with a feeling of moodiness, purposelessness, and social anxiety. However, a third factor known as 'Introvertive Anhedonia' is characterized by schizoid solitariness and a lack of feeling according to Claridge and Beech (1995), whilst Mason et al. (1995) interpret this factor as being characterised by a lack of enjoyment as a result of physical pleasure and social interaction. This is coupled with a withdrawal from both physical and emotional intimacy which is instead replaced with an importance placed on solitude and independence.

The fourth and last factor, 'Impulsive Nonconformity', is considered by Mason et al. (1997) to represent impulse-ridden and disinhibited characteristics as well as self-abusive, violent and reckless behaviours, whilst Friday and Luke (2014) found this sub-scale to be correlated with paranormal belief. It should be noted that Mason et al. (1995) claimed that a more moderate score related to this factor indicates a preference for a non-conforming lifestyle with freedom of choice - this thinking is related to Eysenck's Psychoticism Scale, uncontrollable urges, and hypomania (Mason et al., 1997). It has been argued however that Impulsive Nonconformity may not strictly be an aspect of schizotypy, and is actually more related to borderline thinking.

Well-defined profiles have been highlighted by cluster analyses of these schizotypy dimensions (Goulding, 2004, 2005; Loughland & Williams, 1997; Suhr & Spitznagel, 2001; Williams, 1994; Simmonds, 2003; Simmonds & Holt, 2007) as follows: a) Positive Schizotypes who elicit high scores only on Unusual Experiences; b) Low Schizotypes, who score at a low level on all schizotypy dimensions and are not considered anomaly-prone; c) High Schizotypes, who score highly particularly on Cognitive Disorganisation and Introvertive Anhedonia, and also score highly on Unusual Experiences; and d) Negative Schizotypes, who only score highly on Introvertive Anhedonia. The Positive Schizotypy profile has been referred to as the aforementioned 'Happy' or 'Healthy' Schizotypy as this group is apparently prone to Unusual Experiences, but not susceptible to psychopathology, thus enjoying having higher levels of well-being and mental health than High, Negative and Low Schizotypes (Claridge, 2001; Goulding, 2004; Jackson, 1997; McCreery & Claridge, 1995, 2002).

It would appear that while Unusual Experiences share a significant relationship with a sense of well-being, meaningfulness and optimism (Kennedy et al., 1994; Kennedy & Kanthamani, 1995), low levels of the 'negative' symptoms Introvertive Anhedonia and the 'disorganized' symptom Cognitive Disorganisation are necessary for 'Happy' or 'Healthy' Schizotypy (Goulding, 2004). According to Irwin and Green (1998, 1999), a relationship between schizotypy and anomalous beliefs has been clearly established, and an overlap between such beliefs and psychopathology (Berenbaum et al., 2000; Thalbourne & Delin, 1994) and psychological maladjustment (Irwin, 1991; Thalbourne & French, 1995) has been demonstrated. However, Mehrabian et al. (1997) not only found these beliefs to be unrelated to measures of maladjustment, they suggest these are negatively correlated with measures of psychopathology (Schumacker, 1987).

Further research by Schofield and Claridge (2007) suggested that 'Happy' or 'Healthy' Schizotypes typically evaluate Unusual Experiences as positive, while High Schizotypes tend to describe them as negative. The authors' explanation for this relates to the idea that the Cognitive Disorganisation attributed to these particular schizotypes prevents reassuring belief systems from forming, and it is suggested that 'peculiar' beliefs could act as a protective buffer against stress (Boden & Berenbaum, 2004) and help interpret Unusual Experiences - thus easing or preventing the distress such experiences may cause (Bell et al., 2007). Interestingly, research focusing on childhood concomitants of anomalous beliefs has suggested a second pathway in addition to childhood trauma and demonstrates that childhood fantasy may also lead to such beliefs (Lawrence et al., 1995) which implies imagination plays an essential role in the development of anomalous beliefs.

The mixed literature could be translated to reflect the existence of different types of anomalous beliefs, which may interact in varying ways with the other variables, including those associated with pathology. Possible issues may also lie with the measurements themselves as there has been debates regarding the structure of the Paranormal Belief Scale (PBS) created by Lawrence (1995), which was designed with seven subscales - only three of which (traditional religious belief, precognition, and superstition) actually relate to irrationality (Roig et al., 1998). Additionally, cluster analysis suggests there are different types of 'believer' (Irwin, 1997), namely Traditional Religious Believers, Tentative Believers, Sceptics, and New Age Believers). These four types of believers could relate differently to mental health and pathology. For instance, the literature shows that individuals scoring highly on New Age Beliefs subscale had a far greater relationship with psychopathology than high scores on the Traditional Paranormal Beliefs subscale. It therefore follows that 'belief' in so-called traditional paranormal phenomena may not be associated with pathology as is sometimes suggested.

The choice of analyses themselves should also be considered as whilst a factor analysis groups variables sharing a common variance in a data set, a cluster analysis is intended to group individuals into 'clusters' according to their responses related to a specific set of variables.

Therefore, individuals in the same cluster have a higher similarity to each other than with individuals in other clusters with respect to their patterns of response (Hair et al., 1998). As such, cluster analysis seeks to maximise the similarity or 'homogeneity' of individuals within the clusters, while at the same time maximising the differences or 'heterogeneity' between them. In the process, people are allocated to different groups depending on their responses on the variables of investigation.

Irwin and Green (1998, 1999) found evidence suggesting that the negative dimension of schizotypy could be related to different belief types and that the alignment of both negative and positive schizotypal symptoms may indicate possible psychopathology, and in turn could indicate a relationship with paranormal beliefs rather than when negative or positive traits are found in isolation. This idea was supported by Chapman et al. (1994) who found that although magical ideation seems to be related to psychotic breakdown, the likelihood increases if an individual possesses negative and positive schizotypal traits. Anomalous experiences however have been associated with artistic creativity (Holt, 2007; Kennedy & Kanthamani, 1995), and the idea that the interaction of schizotypal dimensions, rather than the Unusual Experiences subscale alone, may be related to creative functioning (Nettle, 2006) was considered when it was found that the Introvertive Anhedonia subscale was negatively correlated with artistic creativity.

These findings imply that only Happy or Healthy Schizotypy shares a relationship with creative functioning, and it finds support through models which link positive affect with creative functioning (Fredrickson, 2002; Isen, 1999) and the role of determination and focus during the creative process (Eysenck, 1995). Further support was offered by Holt et al. (2008) who suggested that there are two profiles of paranormal believer - the first is associated with High Schizotypy and low well-being levels, and the second is associated with Positive Schizotypy and higher well-being levels. The authors' findings provided further evidence for the notion of the Happy or Healthy schizotype, which supports the clustering approach with regards to exploring personality and belief. The researchers determined that a scale-based approach would likely obscure important differences between the profiles of individuals with high Unusual Experience scores (considered a positive schizotypy trait) and belief in the paranormal. With such a wealth of supporting evidence to suggest that schizotypy may be related to paranormal belief and ability, this trait was measured as a predictor viable in the current study. How Schizotypy will be measured, and whether this scale is meaningful and accurate is detailed in section 2.2.3, but there are important similarities between the Schizotypy subscale 'Unusual Experiences' and the belief in psi measure that may explain possible correlations.

While there are studies which indicate a negative correlation between anomalous experiences and belief in the paranormal, and between unusual experiences and psi activity, (Dagnall et al., 2016; Rabeyron & Watt, 2010), there is some commonality between unusual experiences and belief in psi which may account for a positive correlation between the two variables. Several items on the Unusual Experiences subscale help to measure the participant's perceived ability to sense the presence of others/entities unseen (Mason et al., 1997); for example: 'Have you sometimes sensed an evil presence around you, even though you could not see it?', 'Can some people make you aware of them just by thinking about you?', 'When in the dark do you often see shapes and focus even though there is nothing there?' and 'Do you ever have a sense of vague danger or sudden dread for reasons you do not understand?'.

Williams (1983) found that participants with high levels of belief in psi were more likely to accurately perceive the presence of an unseen person as part of a remote observation experience. This suggests that belief in psi may relate closely to items present on the Unusual Experiences subscale that measure a persons' perceived ability to sense the presence of others who are not physically within one's immediate vicinity, or who are hidden. Therefore, it may be likely that belief in psi and unusual experiences share a positive correlation as, according to Williams' (1983) study, the presence of one of these variables may amplify the other.

Whilst belief in psi 'explores mental processing that occurs on platforms yet to be understood', unusual experiences (according to the Oxford-Liverpool Inventory of Feelings and Experiences) refers to 'perceptual aberrations, magical thinking and hallucinations' (Mason et al., 2005). Although both definitions are unique, both variables share a common theme which is the mental processing by which individuals are able to perceive and interact with the physical world around them (although the unusual experiences subscale focuses more on abnormalities in perception, as opposed to operating on channels yet to be understood). Due to this common theme, belief in psi and unusual experiences may closely relate to each other.

2 STUDY ONE

2.1 Study One Introduction

Having reviewed the literature, it seems that whilst there is support to suggest that self-reports, psychophysiological, and behavioural measures may be valid detectors of remote observation - none are unchallenged, and further investigation is therefore required. Additionally, of the studies which have found evidence to suggest that individuals may be able to detect the stare of others via unconventional senses, many have criticised these approaches for their experimental methodology, or the subsequent analysis of results. Drawing on the existing literature, the research related to Study One described in the following sections was designed based on what has been learned regarding the field of extrasensory detection, whilst incorporating new and original elements to advance what is known about scopaesthesia and its possible predictor variables. The 'sense of being listened to' was included as an exploratory factor, and is untested in previous studies.

2.1.1 Study One Hypotheses.

Having reviewed and considered the literature, the knowledge gained from the vast selection of experiments, surveys, meta-analytic analyses, and documented anecdotal evidence - the following hypotheses were developed. For clarity, the term 'covert surveillance' refers to being watched and/or listened to in a way that the participants could not be aware of via conventional senses.

Formal Hypotheses:

- Participants' correct self-reports will be significantly greater than incorrect self-reports.
- EDA arousal levels will be significantly different in the epochs when participants were under covert surveillance, compared to when they were not.

- The time taken for participants' to accurately complete the Stroop task will be different under covert surveillance, compared to when not under covert surveillance.

Exploratory Hypothesis:

- The ability to detect covert surveillance will be predicted by belief in and experience of anomalous phenomena, temporal lobe lability, and Schizotypy scores, although this is an exploratory hypothesis investigating possible relationships between these factors.

2.2 Methods

2.2.1 Design.

The current study utilised a multivariate mixed experimental design to investigate which psychosocial and neurological factors may predict individuals' ability to correctly report being i) under no surveillance, ii) being listened to, iii) being watched, and iv) being watched and listened to. The individual differences variables were participants' belief in and experience of anomalous phenomena, temporal lobe lability, and levels of schizotypy - all of which were measured prior to the start of the experiment. The repeated measures methodology meant that participants' ability to detect being watched and/or listened to was measured by conscious self-reports scored as correct/incorrect detection, psychophysiological fluctuations in their EDA, and differences in their behaviour and performance whist undertaking a cognitive task. These measurements were taken during periods of randomised surveillance, versus no surveillance which participants could not be aware of via conventional means. Participants' ability to detect being watched and/or listened to was analysed using one-way ANOVA's and paired-samples t-tests, whilst Pearson's correlation analysis and multiple regressions were used to investigate possible trends, patterns and relationships with the individual differences variables.

2.2.2 Participants.

The participants (all aged 18+) were recruited via email campaigns and a dedicated recruitment system (SONA) at a London university, as well as posters (see Appendix A) and advertisements in the press and radio inviting individuals to contribute to a study investigating extrasensory detection. These recruitment initiatives produced 125 respondents, of whom 112 proceeded to take part in the experiment. Due to the recommended sample size ($N > 50 + 8m$) required for multiple regression analysis (Tabachnick & Fidell, 2007), a minimum of 106 participants (see Appendix B) were required for meaningful results. Participants consisted of 66 (59%) females and 46 (41%) males, ranging in age from 18 to 68 years ($M = 31.74$, $SD = 12.99$).

2.2.3 Materials.

The study's data collection was conducted in an observation laboratory comprising of two adjacent rooms - one of which was viewable through a one-way mirror (but not vice versa).

Room 1 (the room in which the participant was observed) featured two desks with a chair, and a computer at each of them. It also had a microphone to amplify and transmit the participant's voice to the adjacent room.

Room 2 (the room in which the observer was situated) featured a chair and a laptop.

Additionally, the study required:

- i) A programme (Qualtrics) to administer the various electronic questionnaires.
- ii) A programme (PsychLab) to run and score the conscious self-report task.
- iii) A computerised random number generator (Random.org) to determine the random sequencing of the trial conditions.
- iv) Three sets of headphones through which music could be played to both participants to control for auditory sensory leakage (and so that the experimenter could monitor conditions).
- v) A coin to determine the order in which the participants were to play the experiments' opposing roles.
- vi) A computerised version of the Stroop Test (Ridley, 1935).

vii) PowerPoint programme to instruct the observer when to watch and/or listen to the participant under surveillance.

viii) A mixing desk to allow the experimenter to control when and what the observer can hear to control for sensory leakage.

ix) Psychophysiological monitoring equipment (a Nexus-4, 4 channel 25-bit ADC from Mind Media) to measure participants' EDA. Please note the following guidelines that were adhered to:

The Society for Psychophysiological Research and Hoc Committee on Electrodermal Measures (Boucsein et al., 2012) outline three different methods for measuring EDA. These include the endosomatic method which does not use the application of an external current and two exosomatic methods, one which uses direct current via electrodes on the skin and one which applies an alternating current instead. Exosomatic recordings using a direct current is seen as the most widely used methodology, but exosomatic recordings using alternating current may offer a reduced risk of error (Boucsein et al., 2012).

Endosomatic Electrodermal Responses

(As described by the Society for Psychophysiological Research and Hoc Committee on Electrodermal Measures)

An electrical potential difference can be measured across the palmar and plantar skin in the absence of any applied voltage. A single electrode is placed on the active site with a reference electrode at a relatively inactive site on the participant's body. The measured potential is usually negative at the palm and electrodermal responses are easily seen in the resulting recordings. The endosomatic electrodermal responses are similar to the more commonly measured exosomatic electrodermal responses, but with a more complex wave form. The measurement of endosomatic electrodermal responses is viewed as the most unobtrusive method (Boucsein et al., 2012).

Exosomatic Measurement with Direct Current

(As described by the Society for Psychophysiological Research and Hoc Committee on Electrodermal Measures)

The measurement of EDA as skin conductance using a direct current, constant voltage methodology with silver-silver chloride (Ag/AgCl) electrodes and an electrolyte of sodium or potassium is considered the most widely used method for measuring EDA (Boucsein et al., 2012). The basic method is to apply a small voltage (such as 0.5V) to two electrodes placed on the intact palmar surface of the skin and include a small resistor in series with the skin. It is recommended that electrodes should be placed on the same side of the body to avoid electrocardiogram artefacts (Boucsein et al., 2012).

However, due to the resistance of the skin, there are concerns that the small series resister is negligible in terms of affecting the current flow in the circuit and can be ignored when measuring current flow. To tackle this, non-polarising silver-silver chloride electrodes is recommended (Boucsein et al., 2012), although there are still existing concerns that, even with non-polarising electrodes, polarisation is still possible. Polarisation is described as, "the counter electromotive force (e.m.f) that is generated at the electrode metal surface" (Boucsein et al., 2012) which causes a voltage opposing the applied voltage. Due to the counter electromotive force, the current is reduced and introduces an error in the recording of skin conductance (SC), which even non-polarising electrodes cannot always tackle.

Exosomatic Measurement with Alternating Current

(As described by the Society for Psychophysiological Research and Hoc Committee on Electrodermal Measures)

Boucsein et al. (2012) argue that the measurement of EDA with alternating current (AC) instead of direct current (DC) has been infrequently used despite showing potential for combating problems associated with using direct current for conductance measurement. Alternating current electrodes virtually eliminates polarization, due to the fact that the continuously changing polarity of an alternating current means that AC polarizes electrode to a much lesser extent than direct current. The direct current system is viewed as non-linear as the measurement is completely dependent on the applied voltage. The alternating current system, on the other hand, counteracts the effects of counter electromotive force due to its reduced polarisation, which means error does not need to be subtracted from the applied voltage when conductance is to be calculated (Boucsein et al., 2012). Therefore, measurement with AC may be more effective as it can reduce the risk of measurement error.

The psychosocial and neurological factors under investigation are listed below, accompanied by the corresponding survey used to measure them.

i) Anomalous Experience and Belief.

This measurement was operationalised by the experience and belief subscales of Gallagher et al.'s (1994) Anomalous Experiences Inventory (AEI) scores (see Appendix C). Thalbourne (2001) found the AEI to have strong convergent validity and reliability (the KR-20 reliability values for the subscales ranged between .64 and .85) due to its intercorrelation with the Australian Sheep-Goat scale (Thalbourne & Delin, 1993) and the Paranormal Belief Scale (Tobacyk, 1988) - both are valid and reliable measures. The version used in the current study consisted of 40 items (out of 70) relating specifically to paranormal experience and belief to measure expectation. All items utilised a 'true' or 'false' response option to questions such as 'at times, I have felt possessed by an outside force', 'I have had waking visions of an event which subsequently occurred', 'I believe that mind can control matter', and 'I believe in life after death'. The first 29 questions related to anomalous experience, whilst the next 11 related to anomalous belief. When analysing the data, the scores were used as a continuous variable, with higher scores indicating a greater level of belief or experience as the measure asks participants to positively endorse their anomalous beliefs and experiences.

ii) Temporal Lobe Lability.

Temporal lobe lability was included as a predictor variable and was measured by Persinger and Makarec's (1993) Complex Partial Epileptic Signs (CPES) and Temporal Lobe Symptoms (TLS) subscales of the Personal Philosophy Inventory scores (see Appendix D). The reliability (internal (alpha) reliability = .70) and validity of these scales has been demonstrated by the substantially higher scores attained by epileptics in comparison to controls, and the greater temporal lobe EEG alpha activity in higher scorers in relation to low scorers (Makarec & Persinger, 1990). The CPS/TLS is a 16-item questionnaire offering respondents 'true' or 'false' options to measure the temporal lobe experiences (TLE) of participants with statements such 'sometimes an event will occur that has special significance for me only', 'when I have a tough decision to make, a sign will be given and I will know what to do', 'I often feel as if things are not real', and 'I have had experiences when I felt as if I were somewhere else'.

iii) Schizotypy.

This predictor variable was be measured by Mason et al.'s (2005) Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE) scores (see Appendix E). The internal consistency of the O-LIFE has been calculated using the alpha coefficient - all four subscales exceeded 0.6, described by Nunnally (1978) as an acceptable coefficient for psychological science (reliability of the scores ranged from .78 to .87). New scales have been correlated with existing scales to calculate concurrent validity and all exceeded 0.9. Burch et al., (1998) have demonstrated the test-retest reliability of the O-LIFE - a 43 item questionnaire with 'yes' or 'no' response options, broken down into the following sub-categories: 1) Unusual Experiences (12 items) containing questions such as 'Have you sometimes sensed an evil presence around you, even though you could not see it?', 2) Cognitive Disorganisation (11 items) containing questions such as 'Do you dread going into a room by yourself where other people have already gathered and are talking?', 3) Introvertive Anhedonia (10 items) containing questions such as 'Are you much too independent to get involved with other people?', 4) Impulsive Nonconformity (10 items) containing questions such as 'Would you like other people to be afraid of you?'

iv) Demographic questions recorded participants' age and gender.

It was expected that each measure would be correlated with the dependent variables (DVs); namely participants' ability to detect being watched and/or listened to via unconventional means. Pearson's correlation analysis and multiple regressions were used to examine whether any such relationships existed.

Measures to determine whether participants detected, or were affected by surveillance (DVs) are listed below, accompanied by a description of the corresponding measurement process.

i) Self-reports (also referred to in the literature as 'guesses') were used to determine whether participants were consciously aware of being watched and/or listened to, and were measured by a programme (PsycLab) designed to accurately establish randomised epochs. During each 30-second trial (the same duration as Lobach and Beirman (2004) used in their experiment), bland text was read aloud (to ensure an audible element to the task), at the end of which the programme asked the participant whether they felt they were being watched or listened to. The participant then indicated their response with a keyboard button push (Y = yes, N = no). The results were subsequently analysed for a relationship between participants' self-reported 'sensed' periods of being watched and/or listened to, and the times during which visual or audio surveillance was actually taking place. To do this, paired-samples t-tests were used to establish whether participants' correctly self-reported being watched and/or listened to significantly more often than they incorrectly self-reported being watched and/or listened to. Pearson's correlation and multiple regression analyses were also used to determine whether participants' self-reports shared a relationship with the individual differences under investigation.

ii) Physiological reactions were determined by electrodermal activity (EDA) which measures the electrical conductance of the participant's skin as an indication of arousal. This forms part of the technique used in traditional lie detector tests, however it is simply a method by which the electrical conductance of the skin is measured. Variations occur depending on the moisture of the skin which increases and decreases with the amount of sweat excreted by the individual under observation. The sympathetic nervous system is responsible for this reaction (Martini et al., 2003), and so EDA was used as an indicator of physiological or psychological arousal. Significantly increased or decreased EDA readings (averaged across 2 x 30-second epochs) were therefore taken to indicate that the participant was showing physiological signs of detecting covert surveillance. The resulting data was analysed using a one-way ANOVA and paired-samples t-tests to investigate significant differences between participants' EDA in each of the surveillance conditions compared with control (no surveillance). To investigate whether participants' EDA shared a relationship with the individual differences under investigation, a Pearson's correlation analysis and a multiple regression analysis were conducted.

iii) Significant changes in participants' task performance were used to measure whether remote surveillance had been detected. To test the possible effect of remote social facilitation, each participant performed a computerised version of the Stroop Test (Ridley, 1935) under four different surveillance conditions (a) not under any surveillance (control group), b) being listened to, c) being watched, and d) being watched and listened to. To introduce an audible element to the task, participants had to say the colour of the key they pressed to indicate their response aloud. The results were subsequently analysed via a one-way ANOVA and paired-samples t-tests for whether the surveillance conditions significantly affected the time participants took to complete the Stroop Test, compared with when they were not under surveillance. Pearson's correlation and multiple regression analyses were conducted to determine whether participants' self-reports shared a relationship with the individual differences under investigation.

2.2.4 A Brief Overview of the Stroop Test

The Stroop Effect (Ridley, 1935) is a cognitive task in which individuals are asked to choose the colour a word is written in, rather than the colour it actually says. For example, if the word 'yellow' is written in red, the participant must say or choose "red". This is more difficult for most people than it sounds as the actual words themselves have a strong influence over an individuals' ability to say of choose the colour in which it is written. The interference between the differing information confuses the brain, as what the words say competes with the colour of the words, thus causing interpretation difficulties.

One of the explanations for this difficulty is that people are so used to processing word meaning, while ignoring the physical features of words that it has become a learned response (Ridley, 1935). The Stroop task requires individuals to do something that they have never learned, and which is the opposite of what they would normally do. Therefore, the processing of one dimension requires much more attention than does processing of the other dimension. Therefore, naming the ink colour draws more heavily on attentional resources than does reading the irrelevant word. Moreover, reading the word is seen as obligatory, whereas naming the ink colour is not. Presumably, this imbalance derives from our extensive history of reading words as opposed to naming ink colours (Kahneman & Chajczyk, 1983). According to this view, the asymmetry that is the fundamental characteristic of the Stroop task must occur. Put another way, words are read very automatically, whilst colours require considerably more attention to be named correctly.

Because of the extensive practice that naturally occurs with word reading in the standard colour-word version of the Stroop task, there have been numerous variations such as when Dunbar and MacLeod (1984) rotated the colour words (e.g., upside down and backward) to create a version of the task in which practice could be controlled and observed from the outset. The researchers found that their participants' reading of the rotated words was dramatically slower than reading normal words, and that whilst word reading was still far slower than colour naming, incongruent words still interfered with colour naming. As such, the slower process was affecting the faster one, which is a finding inconsistent with the relative speed of processing account.

Kahneman and Chajczyk (1983) called into question the automaticity view with their demonstration of 'dilution' of interference. Using a modification of the Stroop task in which the colour word appeared below a colour patch, they observed that an additional word in the display actually reduced the interference when naming the colour patch. Apparently, another dimension of the stimulus display could split attention, thereby lessening the impact of the colour word on naming of the colour patch. This dilution is inconsistent with the idea of automatic reading, as attentional allocation should not affect the amount of interference observed if the process is automatic.

Differing variations have tested for response modality (pressing a button or reading the word aloud), hue variation, which colours are included, how many colours are used, the proportion of incongruent to congruent trials, whether congruent or incongruent trials are 'blocked', participant characteristics, and age and gender (McLeod, 1991). However, for the current experiment the researcher decided on using the classic Stroop Test as its effects are so well supported that it has been used in research with clinical groups to test for disabilities such as ADHD (Banks, 2017).

2.2.5 Focus Group and Pilot Studies.

To ensure the study met with the ethical standards of the British Psychological Society (BPS) as well as the University of Greenwich where the study was taking place, a focus group was held with eight group members. Half were university students, and half were interested members of the public. This mix was deliberate to try and replicate the intended study sample as it was predicted that due to the proximity of the laboratory, students would form around 50% of the participant sample - however, an entirely student sample was actively avoided to ensure the results could be generalised to the wider public.

Ensuring varied focus group members was also important because the main topic of the focus group was to measure participants' likely reaction to being covertly watched and listened to. Whilst the participants would be made aware that this would take place, they would not know when. It was thought that whilst university students - particularly those studying psychology may be more sympathetic to the necessity for this element of the research, it was deemed necessary that members of the public without knowledge of psychological research issues be included in the focus group. Despite this concern, no discernible or salient difference was found between the student participants and members of the general public in their attitude towards covert surveillance within the context of the study.

During the course of approximately 20 minutes, the focus group leader enquired as to the theoretical reaction of those within the group regarding covert surveillance, and in every instance the participants understood the necessity for this and its role in enhancing the understanding of remote observation detection. It was also clearly explained that every participant would have the opportunity to abort their participation should they experience any discomfort or negative reactions, and all involved in the focus group stated that they personally had no ethical concerns with the study and that they would not expect others to either.

At the end of the session, the group was given the opportunity to voice any concerns they had with any other elements of the research as the general experimental design had been described to them in detail. Concerned that perhaps some members of the group were simply uncomfortable raising issues directly or in front of others, the researcher gave them the opportunity to do so via email within one week of the focus group taking place. No such emails were received, and only after this seven-day period expired did the first pilot study proceed.

2.2.6 Pilot Study One (Lab Layout and Controlling for Sensory Leakage).

For the first pilot study, expecting there to be methodological issues the experimenter took the role of both the researcher and the participant within the actual lab that was to be used to enable him to identify initial problems of which three became immediately obvious. The first concern was that the one-way mirror was slightly see-through, however this was easily resolved with an adjustment of the lighting. By ensuring there was significantly less light in the observation lab than there was in the adjacent room, the one-way mirror became reflective enough that there was no way for the participant under surveillance to see whether they were being watched or not. It should be noted that every participant had their back to the one-way mirror during all testing periods. This served not only to replicate the position of participants described in the majority of previous studies, but also to further protect against sensory leakage explaining possible positive results.

The room was also not laid out in a way that was conducive to continuing the work of previous researchers who have typically positioned their participant under observation directly in front of, but facing away from the observer. With the help of the lab technicians who rearranged the computers this was resolved so that the experimenter's set up closely resembled the positioning used by the majority of the studies on which the current study was based. This also provided another element of protection against sensory leakage as the participant under observation would then face a wall with the surveillance participant directly behind them, making them theoretically impossible to see. However even movement and shadows which could have been detected in the participant's peripheral vision were eradicated by the one-way mirror which had already been tested to be effective.

The third problem the experimenter encountered as the observer/listener was more challenging to overcome as it became clear that if a recorded message was played in the observation half of the laboratory (to represent what would eventually be the voice of an actual participant when the experiment took place for real), it could not be reliably heard. When the recorded message was played at a volume the experimenter and his colleagues estimated that participants could reasonably be expected to maintain for minutes at a time, the voice was only just audible. It was predicted that some participants may not be comfortable talking at high volume, and some may become increasingly quieter if they forget the importance of being heard - additionally, the person observing may either create accidental noise or not have perfect hearing, and so it was deemed necessary to overcome the problem with technology. This methodological addition would not only assist with ensuring valuable data could not be lost or may be unreliable - it also helped the study's validity as in reality, people under surveillance are unlikely to be shouting or speaking unusually loudly.

This methodological problem was overcome via the use of a microphone set up in room one, which was connected directly to headphones in the adjacent room. This not only removed the need for the participant to talk at an unnatural volume, but also meant that it could be adjusted every time to ensure the observer/listener could hear the words being spoken regardless of their individual hearing ability as the volume could be turned up if necessary by the experimenter. The introduction of the microphone and earphones to the experiment also presented a way to control what had been an area of concern since its inception - namely how to dictate exactly what and when the observer/listener can hear.

Whilst being able to listen to the participant under surveillance is essential for the observer/listener, being able to guarantee they are *unable* to listen at certain points was also essential if a difference in response was to be determined by when they are heard, and when they are not heard. By playing bland and unemotive music through the headphones during the periods when the participant under surveillance should not be heard at a volume which makes hearing their voice impossible, the experimenter was effectively able to turn the observer/listener's hearing on and off as required. The experimenter considered that the music being played to the participants could present an attentional distraction from the task at hand. In the case of the observed participant, the music could have limited their focus on the relatively simple self-reporting task, and possibly more so during the complicated cognitive task (the Stroop Test). Similarly, as the music was chosen to not be arousing in an attempt to *avoid* such distraction, it could have had the unintended effect of relaxing the participants. If this were the case, it could have affected all protocols - not least the EDA measure as the results of this protocol relied on the differences in participants' physiological states. Indeed, Lobach and Beirman (2004) themselves considered whether their use of music may have resulted in a calming effect which could have over-ridden any activating effect of remote observation. These researchers never investigated whether this was the case unfortunately - however when the current experimenters' pilot tests showed that without the audio, the observed participant could notice sensory cues emanating from the observers' movements in the adjacent laboratory, it was decided that audio played through headphones was unavoidable. Audio was also regarded as necessary for the observer participant despite the possible issues it may cause, as pilot tests demonstrated that there was no better way for the experimenter to control when the observer participants could (and could not) listen to the observed participant with the technology, apparatus and room layout available for the research.

2.2.7 Pilot Study Two (How to Instruct Participants).

With the audio issues resolved, and the rest of the experiment seemingly working well when the experimenter tried both roles - a second pilot study was run with a fellow researcher so that it was conducted in a way more representative of how it would take place in reality (with the role of the observed and the observer taking place simultaneously). Whilst this appeared to reveal no issues when this colleague played the role of the participant under surveillance, it did highlight a problem when they took the role of the observer. When the primary experimenter played this role and simply tested whether the line of sight made effective observation possible, the one-way mirror was effective, and speech could be clearly heard from the adjacent room there were no obvious issues; however this left another problem to overcome.

Whilst due to the introduction of the headphones following the first pilot study the experimenter could now control what the observer participant heard, it did mean that they were constantly listening to either the voice of the observed participant, or the music intended to ensure they could not. In either instance, they would not be able to hear the experimenter's instructions as to when they should observe their counterpart, and when they should stop doing so. These instructions were originally intended to be given verbally, but the need to automate them became obvious - however how to do so was not so clear.

Using PowerPoint to issue timed instructions provided a solution when the observer participant was facing away from the participant in the adjacent room by merely placing a laptop running this programme within eye line. The difficulty was in how the observer participant would know to turn around whilst they were observing as it was impossible to place another laptop running the same programme issuing instructions directly in front of them without blocking their view of the very person they were meant to be staring at. Eventually this conundrum was overcome by using the fact that the observer was looking through a mirror.

By making every other instruction within the programme a mirror image which could be read perfectly as a reflection, it was possible to issue the instructions via a single carefully placed laptop which could still be positioned behind the observer participant. When they were facing away from their counterpart, the observer could read the instruction to turn around exactly as they could before; however, once facing them, they could still see the reflection of the same screen and read the reversed message to know when to change position. Tests of this with eight individuals showed that this new system was easy to read and understand, and all turned towards and away from the observed participant at the correct time when tested. Additionally, they all reported that they were able to stare intently at the person in the room next door without distraction from the reflection. The experimenter used the same instructions issued via the laptop to control when the observer participant could hear the observed participant and when they could not via headphones as described in section 2.2.6. Please see the experiment diagram (Figure 1) for the laboratory layout, and the participants' (and experimenter's) position within it.

This also resolved another important theoretical methodological issue. If scopathesia is a genuine phenomenon, and if it can affect a person's behaviour as hypothesised - then it follows that the experimenter watching the observer participant may result in an artefact. Firstly, being observed by the experimenter may affect the participant who is supposed to be observing along with their ability to follow instructions correctly. Secondly, perhaps a 'doubling-up' of observation may occur and this could affect the intensity of the stare and therefore influence detection. By taking away the experimenter's need to watch the participants, this potentially important issue was dealt with.

2.2.8 Pilot Study Three (Further Sensory Leakage and Sequencing Issues).

With the experiment significantly closer to being ready, its methodology was again tested in a further pilot study that incorporated all of the changes and improvements so far. This time, the experimenter did not take part, but studied his colleagues playing the role of both the observer and the observed participants thinking that a fresh perspective may reveal additional flaws. It did. Whilst it was impossible for the observed participant to see whether the observer participant was watching or not, it *was* possible in some instances for the participant under surveillance to hear their counterpart turning around. Whilst even if the observed participant did detect this, it would not tell them which way their counterpart was facing - it would tell them they had turned (the observer is not necessarily instructed to turn every 30 seconds, with minutes sometimes passing with no change in position). For this reason, it was determined that this possible issue of sensory leakage should be dealt with, which was achieved via another set of headphones worn by the participant under surveillance. Bland, unemotive music was played to them at a volume which made the already feint sound from the adjacent room inaudible.

Further changes were made to the initial experimental design, not as a direct result of the pilot studies, but due to the completion of the literature review which highlighted issues with possible sequence guessing. This potential issue was also raised during conversations following a presentation by the experimenter at an international conference where he spoke at length regarding the methodology and the intended investigation. Interested members of the audience were also concerned that by chance, participants' estimations of the likely ordering of conditions may accidentally coincide with the actual sequence, even though they were created by a random number generator. To combat this possible artefact, three more alternative programmes were created (again via the use of a random number generator) to instruct the observer participant when to observe and/or listen. Every pair of participants were then randomly assigned one of these four sequences (see Appendix F) to further decrease the possibility of sequence guessing accounting for any accuracy found.

2.2.9 Procedure.

Due to the complexity and ethical considerations necessary for research that includes covert surveillance, in addition to adhering to the standard British Psychological Society (BPS) Code of Ethics and Conduct of briefing participants, inviting and answering their questions, obtaining consent (see Appendix G), ensuring confidentially, informing them that that they can withdraw themselves and their data at any point, and debriefing them fully (see Appendix H); a focus group was held prior to the experiment taking place (see section 2.2.5).

Participants could apply in pairs, or as individuals to be paired with another participant by the experimenter. It should be noted that in instances when one participant did not arrive, the experimenter's intention was to take their place as the observer so as to not waste participants' time or the data they could provide. Such an event would always have been recorded so that whether this affected the results could be analysed, however this situation never occurred.

Participants were issued a specific time and place at which the study would be conducted, and upon arrival both participants were offered a seat at a computer (each participant was in a separate room to ensure confidentiality). They were instructed to read and complete the consent form issued to them, and were offered the chance to ask any questions they may have. All the individuals who applied subsequently consented to take part. They were then invited to complete four questionnaires; the AEI, the O-LIFE, the CPES/TLS, and a short demographic questionnaire. All questionnaires were completed online using the Qualtrics survey programme to allow the participant to enter their responses directly into the database to eradicate possible errors in transfer. The experimenter was present, but was never directly observing to avoid participants' performance anxiety, or demand characteristics.

Once both participants had completed the questionnaires, which of them was to first play the role of the 'observed' participant (the participant under surveillance), and who was to be the 'observer' participant (the participant watching and/or listening) was decided by a coin flip. All participants played both roles. The order of the protocols (self-reports, EDA, task performance) was decided by a random number generator to avoid artefacts such as practice effects.

2.2.10 Study One, Protocol One (Measured by Self-Reports).

The observed participant remained in the observation lab and sat at a computer with the PsycLab software pre-loaded. They were instructed to read aloud the text that would appear on the monitor in front of them for 20-seconds until the programme asked via the screen whether they felt they had been watched, at which time they were to indicate their answer on the keypad by pressing 'Y' for 'yes' and 'N' for 'no'. Similarly, the programme then asked them to register whether they felt they had been listened to, and the participant answered this second question in the same way. It was explained to the participant that this would happen eight times, and that during any of these trials the participant could be a) under no surveillance at all, b) watched, c) listened to, d) watched and listened to. It was made clear that the observer participant's task would last four consecutive minutes, during which time any of the above conditions may occur. It was explained that these periods of observation may take place once, multiple times, or not at all. Further explanation was given if required.

The observed participant was instructed to start reading the words aloud immediately upon the programme starting (which was delayed to ensure the experimenter had time to join the observer participant in the adjacent room). During this time, the observed participant was instructed to relax and to listen to the music played through headphones and it was emphasised that they must not remove them until the programme indicated they should do so (to ensure that audio based sensory leakage was not possible). This brief period of time was used to familiarise the observer participant with the adjacent observation area and the equipment used to control when (and for how long) they would watch and/or listen to the observed participant. This equipment consisted of headphones through which the experimenter could control whether the observer participant could hear the same music as the observed participant, or them reading aloud. It was made clear that during this time the observer participant should intently listen to the words being spoken until the headphones returned to emitting music once gain (to ensure the observed participant was only audible during the designated periods). It was also explained that visual instructions issued via a computer would dictate when the observer participant should turn and stare at the observed participant through the one-way mirror, and when to look in the opposite direction.

The experimenter decided on the duration of each condition based on previous research, as well as participants' likely attention span for the task (informed by the pilot study). Each of the four conditions (30-seconds in length including time for the participants to consider and respond to the questions asked) occurred twice - so the task lasted four minutes in total. The results were analysed via paired-samples t-tests to investigate whether participants' correct self-reports of when they were being watched and/or listened to were higher than their incorrect reports, as it was suggested by Atkinson (2005) that remote observation trials should be analysed by comparing the total hit rate and the total false alarm rate. Pearson's correlation analysis and multiple regressions were then used to examine possible patterns and relationships with the individual differences variables under investigation.

2.2.11 Study One, Protocol Two (Measured by EDA)

For this element of the experiment, the participant who was to be watched and listened to first was fitted with EDA equipment according to international guidelines (see section 2.2.3) so that the electrical conductance of their skin could be measured to indicate physiological arousal. The pads used to take the readings were carefully placed and strapped onto the tip of the participant's index finger, and their middle finger. A base level 'resting' reading was taken to ensure the appliance was in working order and had been fitted correctly. It was explained to the participant to that they must remain still so as to not adversely affect the EDA readings, and that they must read aloud from bland text from a printout situated in front of them so that they could be listened to by the observer participant. They were informed that this task would last for approximately five consecutive minutes (the first minute was to return to a resting state), during which time they would be a) completely free from any surveillance, b) listened to, c) watched, or d) watched and listened to). It was made clear that these periods of surveillance may occur once, multiple times, and in any order. Further explanation was given if required, and EDA adjustments made if necessary.

The observed participant was instructed to assume a comfortable position and begin reading the words aloud, but to relax, breathe normally and stay as still as possible (to return arousal levels to a resting state, and so that subsequent changes in their EDA could be attributed to remote surveillance detection rather than physical movement) until instructed otherwise. The experimenter then left the room to familiarise the observer participant with their task, which was similar to that described in section 2.2.10 whereby their ability to hear either music or their counterpart reading was controlled by the experimenter, and the 'stare' and 'do not stare' instructions were issued by the computer programme. The same four conditions (no surveillance, listened to, watched, and watched and listened) occurred twice for 30-seconds each meaning the observer participant's task lasted four minutes. The order of these surveillance conditions were again randomised by the number generator.

To determine whether each individuals' EDA differed significantly from their own baseline in the experimental conditions, significant deviation was gauged against participants' control condition, i.e. their EDA when under no surveillance because an individuals' EDA score needed to be considered relative to their own baseline. For this reason, a one-way ANOVA and paired-samples t-tests were conducted to examine whether there was a significant difference in participants' EDA readings when they were under no surveillance, with when they were listened to, watched, or watched and listened to. Pearson's correlation analysis and multiple regressions were subsequently used to investigate possible relationships with the individual differences variables.

2.2.12 *Study One, Protocol Three (Measured by Task Performance)*

Similarly to in the self-report and EDA protocols, for the task performance element of the experiment the observed participant sat at a computer in the observation lab - but this time with a pre-loaded computerised version of the Stroop test (Ridley, 1935) on the screen in front of them. The basic principles of the Stroop test were explained to them, and it was made clear that they should press the letter on the keypad which corresponded with the colour that each word which appeared in front of them was written in, rather than what it actually said. For instance, if a word appeared which read 'yellow', but was written in 'blue' - the participant should press the 'b' key. Likewise, if a word appeared which read 'green', but was written in 'red' - the participant should press the 'r' key. It was also made clear to them that they should read the words aloud as they appeared (so that they could be heard by the observer participant).

A short practice session was held in the experimenter's presence to ensure the task was understood; however, no problems arose with participants' understanding. It was explained to each participant that they would take the Stroop test four times, and that during any of these trials they could be a) under no surveillance at all, b) listened to, c) watched, d) watched and listened to. The participant was then instructed to wear the headphones to ensure they could not hear any movements or conversations in the adjacent room.

The observer participants' job was simpler in this protocol, and for each trial they either had to a) look away from the one-way mirror and listen to music (to ensure they could neither see nor hear the observed participant), b) look away whilst listening to the words being read aloud (so they could only listen to the observed participant), c) watch the participant intently whilst listening to the music (so they could stare at the observed participant, but not hear them), or d) watch the participant intently and listen to the words they are saying (so that they could both see and hear the observed participant). The observer participant knew which of these conditions to adhere to via prompt cards held up by the experimenter. The duration of each trial depended in the speed with which the observed participant completed each Stroop test, but each trial was in the region of four minutes. The process was repeated until the task had been completed under all four conditions, the sequencing of which was randomised by a number generator. The computerised task automatically recorded the accuracy and speed of the observed participants' responses so that their reaction times (as no participants made any errors) could later be analysed for relationships with the surveillance conditions via a one-way ANOVA and paired-samples t-tests, whilst patterns and trends with the individual differences variables were investigated via Pearson's correlation analysis and multiple regressions.

Upon completion of all three protocols, participants' roles were then switched so that the observed participant became the observer. This was designed so that the results could not be attributed (positively or negatively) to the possible psi ability (or inability) of the individual watching and/or listening to the observed participant. However, in cases where only one participant arrived, the duties of the observer would have been performed by the experimenter, and this would have been noted to be accounted for during analysis. The below chart summarises how the experiment ran.

Table 2

The Sequence of Events for Both Participants

		Measurement of surveillance detection (randomised)		
Participant 1	Answers online questionnaires	Self-report (4 minutes)	EDA (5 minutes)	Task performance (4 x 4 minutes)
Participant 2	Surveillance sequence training	Provide surveillance	Provide surveillance	Provide surveillance

The debrief form was then issued, and participants were reminded of their freedom to withdraw their data, and that they should seek support if they found any elements of the study unsettling. Whilst the importance of not relaying details of the experiment to anyone else was specifically mentioned in the debrief form, however this was also stressed verbally before the participants left the laboratory due to the fragility of this element of the experiment, and its dependency on future participants' unfamiliarity with the methodology.

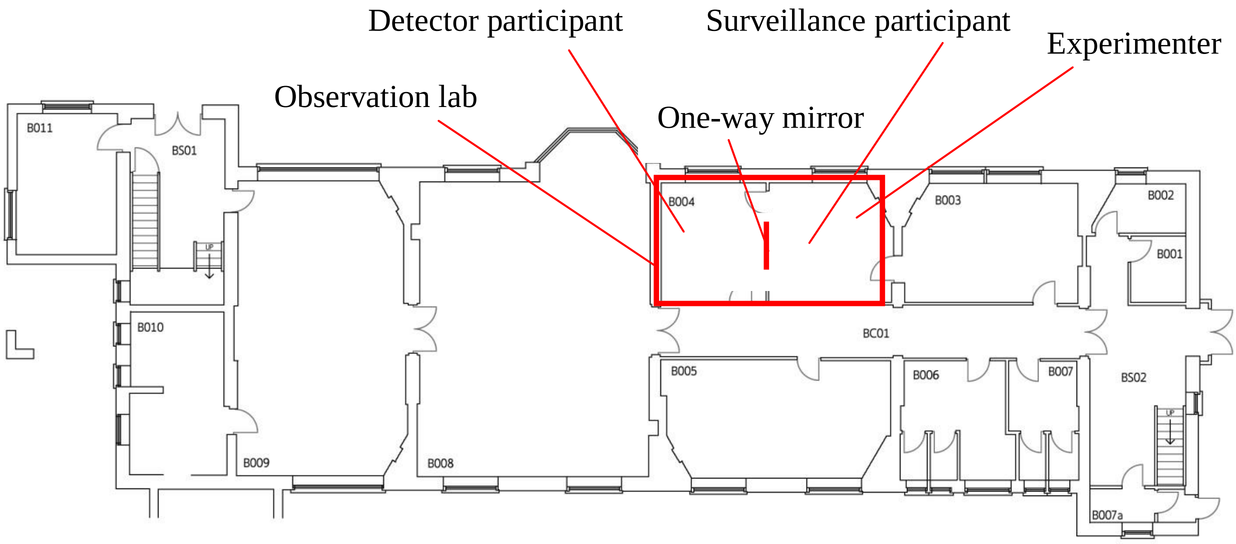


Figure 1 Layout of the adjacent laboratories and key placements for reference.

Note. The above plan (of the ground floor of the University of Greenwich Psychology Department) shows the layout of the labs in which the experiment took place, as well as the positions of the experimenter and participants.

2.3 Study One Results (Detection Measured by Self-Reports)

2.3.1 Treatment of the Imperfect Self-Report Data.

There were four missing data points within the self-report results. These could have been due to indecision on the participants' part, or by not responding when prompted by the computer programme within the designated time (five seconds). To ensure a complete set of results, these missing data points were entered using the equivalent response for that participant e.g. if the participant had responded 'yes' when asked whether they felt they were being watched on one occasion, but omitted a response when asked the same question under the same condition, the experimenter assumed the same response. Out of 1,792 possible self-report responses, only four were missed and no more than two occurred within the same participants' data. It was decided that whilst adding these data points may create marginally more noise, their omission would have excessively complicated the already complex data. As half of the missing data points added were in the positive direction, and the other half were negative - the changes effectively amount to zero, whilst avoiding missing data and recalculating all the probabilities. The only other alternative was random allocation, which would have amounted to the same outcome. Additionally, the experimenter felt that as the missing data points represented just 0.22% of participant responses - this was unlikely to affect the results in a systematic way, and that dismissing otherwise valuable participant data was wasteful.

2.3.2 Analysing the Self-Report Scores for Possible Bias.

Sheldrake (1998, 1999, 2000, 2001, 2005b) highlighted that response biases may play a role in apparent scopaesthesia results, and noted that a bias to answer positively more often than negatively exists (2005b). Atkinson (2005) also made the argument that response bias should be accounted for in future studies, and initial analysis of the self-report scores did indeed indicate a general tendency for participants to report the feeling of being observed in trials (see Appendix I), with 950 positive observation responses overall out of a possible 1792. This was equivalent to a 53% response rate, which was significantly higher than mean chance expectation (50%) where $t(111) = 2.08$, $p = 0.04$ (two-tailed). There was also a greater tendency to report being watched than being listened to, however this difference was not significant $t(111) = 1.32$, $p = 0.19$ (two-tailed), and the reported observations in each mode exhibited a small significant positive correlation between reporting being watched and reporting being listened to r(112) = .34, *p* = .0002, indicating a fairly consistent within-subjects tendency to report being observed regardless of the speculated mode of observation.

2.3.3 Analysing the Self-Report Scores Using Paired T-Tests.

Paired-samples t-tests (see Appendix J) were conducted to investigate whether participants' correct self-reports differed significantly from their incorrect self-reports. The data did not pass the assumptions required (see Appendix K) for one-sample and paired samples t-tests. The Kolmogorov-Smirnov test of normality shows the data to not be normally distributed, however the Shapiro-Wilk test of normality is more powerful and suggests that it is significant (just). Additionally, closer inspection of the histogram shows that the data resembles a normal distribution. Despite the imperfect distribution of the data, t-tests remain a valid analysis due to the study's large sample size and the robustness associated with t-tests (Lumley et al., 2002).

Table 3

*Signal Detection Rates for the 'Listened to' Mode of Response for all Trials (Nt = 2) for all Participants (N = 112) for all of the Conditions **

Watched	True	True %	False	False %
Positive	238 (hit)	53.1%	225 (false alarm)	50.2%
Negative	223 (correct rejection)	49.8%	210 (miss)	46.8%
Mean		51.4%		48.6%

*[total number of trials across all conditions and all participants = 896]

While trends were in the predicted direction, a paired-samples t-test comparing correct Listened To self-report scores (Listened To and Watched and Listened To conditions), *M* = 2.13, *SD* = .98, with incorrect Listened To self-report scores (Watched and No Surveillance conditions), *M* = 2.01, *SD* = 1.11, was not significant; *t*(111) = .94, *p* = .35 (one-tailed), *d* = . 11 (-0.15, 0.38).

Table 4

*Signal Detection Rates for the 'Watched' Mode of Response for all Trials (Nt = 2) for all Participants (N = 112) for all of the Conditions **

Watched	True	True %	False	False %
Positive	244(hit)	54.5%	243 (false alarm)	54.2%
Negative	205(correct rejection)	45.7%	204 (miss)	45.5%
Mean		50.1%		49.9%

*[total number of trials across all conditions and all participants = 896]

While trends were in the predicted direction, scores were almost exactly at chance, and a paired-samples t-test comparing correct Watched self-report scores (Watched and Watched and Listened to conditions), *M* = 2.18, *SD* = .96, with incorrect Watched self-report scores (Listen and No Surveillance conditions), *M* = 2.17, *SD* = 1.0, was not significant; *t*(111) = .07, *p* = .47 (one-tailed), *d* = .01 (-0.25, 0.27).

Table 5

*Signal Detection Rates for the 'Both' Mode of Response for all Trials (Nt = 2) for all Participants (N = 112) for all of the Conditions **

	True	True %	False	False %
Positive** (MCE = 25%)	72 (hit)	32.1%	205 (false alarm)	30.5%
Negative*** (MCE = 25%)	62 (correct rejection)	27.7%	163 (miss)	24.3%
Mean		29.9%		27.4%

*[total number of trials across all conditions and all participants = 896 (224 for true, 672 for false]

** True/False Positives are those trials where participants responded that they had been both Watched and Listened to

*** True/False Negatives are those trials where participants responded that they had not been either Watched or Listened to

While trends were in the predicted direction a paired-samples t-test comparing correct combined both Watched and Listened To self-report scores ('Both' condition), M = .64, SD = .66, with incorrect combined both Watched and Listened To self-report scores (None condition), M = .61, SD = .38, was not significant, $t(111)$ = .49, p = .31 (one-tailed), d = .06 (-0.21, 0.32).

Table 6

*Signal Detection Rates (Red = Above Chance, Blue = Below Chance) for all Responses Combined, for all Trials (Nt = 2) for all Participants (N = 112) for all of the Four Conditions **

	True	True %	False	False %
Positive** *N* = 1344	616 (hit)	45.83%	836 (false alarm)	61.45%
MCE		41.67%		58.33%
Negative*** *N* = 1344	562 (correct rejection)	41.81%	782 (miss)	58.18%
MCE		41.67%		58.33%
Total		43.82%		60.19%
% above MCE		2.15%		1.86%

* Total number of trials across all conditions and all participants = 2688 (1344 for positive, 1344 for negative

** True/False Positives are those trials where participants responded that they had been under surveillance.

*** True/False Negatives are those trials where participants responded that they had not been under surveillance.

While trends were in the predicted direction a paired-samples t-test comparing correct combined self-report scores for all four conditions (None, Listen, Watched & Both conditions), M = 5.51, SD = 2.06, with incorrect combined self-report scores for all four conditions (None, Listen, Watched & Both conditions), M = 5.18, SD = 1.91, was not significant, $t(111)$ = 1.15, p = 0.125 (one-tailed), d = .17 (-0.10, 0.43).

2.3.4 Transforming the Self-Report Scores into Ratio Scores.

In regards to the exploration of correlations with individual differences, due to the reporting bias evident in participants' responses (see Appendix I), a self-report ratio score was calculated to account for this. This new self-report ratio score was created by taking each of the self-report scores for reporting accuracy (e.g., reporting being listened to when they had been listened to) for each condition (no surveillance, listened to, watched, watched and listened to) and dividing each of these totals by the self-report total of positive responses overall so that the ratio score is relative to the individual participants' overall reporting of surveillance detection to correct for the general tendency to over-report surveillance detection.

Table 7

Recalculated Self-Report Ratio Descriptive Statistics

Self-report ratio conditions	Mean	Std. Deviation
No surveillance (*N* = 2)	0.15	0.09
Listened to (*N* = 4)	0.34	0.08
Watched (*N* = 4)	0.34	0.10
Watched and listened to (*N* = 2)	0.17	0.05

2.3.5 Correlation Analysis of Individual Differences (Self-Report Ratio Scores).

Participants' self-report ratio scores related to Study One were analysed for relationships with each of the individual differences (see Appendix L for descriptive statistics) under investigation to examine any meaningful patterns or trends. The below are those found to be statistically significant using Pearson's correlation co-efficient analysis (see Appendix M), however none remained significant after a Bonferroni correction for multiple analyses was applied.

Table 8

Correlations Between Participants' Individual Differences Variables and Self-Report Ratios

N = 112	No surveillance (*M* = 0.07, *SD* = 0.76)	Listened to (*M* = 0.25, *SD* = 0.11)	Watched (*M* = 0.26, *SD* = 0.13)	Watched and listened to (*M* = 0.07, *SD* = 0.07)
AEI Experience (*M* = 9.11, *SD* = 5.37)	.112	-.090	-.047	.094
AEI Special Experience (*M* = 10.42, *SD* = 5.568)	.095	-.103	-.038	.111
AEI Belief (*M* = 5.02, *SD* = 2.92)	.151	-.130	-.076	.064
AEI Special Belief (*M* = 6.43, *SD* = 3.368)	.111	-.141	-.062	.073
Unusual Experiences (*M* = 4.49, *SD* = 2.71)	.209*	-.086	-.042	.245**
Cognitive Disorganisation (*M* = 4.65, *SD* = 2.88)	.125	-.121	.034	.131
Introvertive Anhedonia (*M* = 2.95, *SD* = 2.42)	.064	.280**	-.106	.039
Impulsive Non-Conformity (*M* = 4.01, *SD* = 5.37)	.046	-.001	-.057	.070
CPES (*M* = 4.49, *SD* = 3.285)	.220*	-0.1254	-.007	.220*

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

NOTE: None of the significant correlations remained so after the Bonferroni correction (alpha = .0055).

2.3.6 Regression Analysis of Individual Differences (Self-Report Ratio Scores).

The association of individual differences to self-report ratio scores were explored via four multiple regressions, with the observed variable conditions; (none, listened, watched, both) as outcome variables and the nine individual differences measures (AEI experience [x2], AEI belief [x2], Unusual Experiences, Cognitive Disorganisation, Introvertive Anhedonia, Impulsive Nonconformity, and CPES) as predictor variables (see Appendix L for descriptive statistics). All assumptions for multiple regression (linearity, non-multicollinearity, residual independence and normal distribution, homoscedasticity, absence of influential cases, and sufficient power) were met. Of the four multiple regressions (see Appendix N) none returned significant models.

2.4 Study One Results (Detection Measured by EDA)

2.4.1 Analysing the Results of the EDA Measure.

To determine whether changes in participants' EDA can be used to measure surveillance detection, participants' mean EDA values (in micro Siemens) under the four surveillance conditions (no surveillance, listened to, watched, watched and listened to) were measured. Each of these four conditions lasted 30 seconds and occurred twice for each participant. The descriptive statistics (see Appendix O) for the EDA measure can be seen in Table 9 below.

Table 9

Mean EDA Descriptive Statistics

Mean EDA score	Mean micro Siemens of both 30 second epochs combined	Std. Deviation
No surveillance	6.54	2.98
Listened to	6.58	2.95
Watched	6.48	3.17
Watched and listened to	6.39	2.86

2.4.2 Analysing the EDA Scores Using a One-Way ANOVA.

A one-way repeated measures ANOVA was conducted to determine whether there was a statistically significant difference in EDA means across conditions (none, watched, listened to and both). Mauchly's tests of sphericity indicated that the assumption of sphericity was violated, X^2 (5) = 30.29, p = .00001. Epsilon was 0.83 as calculated according to Greenhouse and Geiser (1959), and was used to correct the repeated measures ANOVA (see Appendix P). There were no statistically significant differences in EDA means across conditions F(2.49, 276.77) = 1.72, p = .17, partial w^2 = .015.

2.4.3 Transforming the Initial EDA Values into EDA_d Values.

As it could be considered meaningless to look at general magnitude without factoring in intra-individual differences from a no observation baseline, new EDA difference figures (EDA_d) were calculated. These were created by deducting each participants' 'no surveillance' value from their surveillance (listened to, watched, watched and listened to) values to take into account that the EDA means do not consider individual variances in EDA readings. The mean and standard deviation for these recalculated EDA_d values can be seen in the below table (see Appendix Q).

Table 10

EDA_d Descriptive Statistics

Recalculated EDA_d values	Mean	Std. Deviation
Listened to	0.004	0.13
Watched	0.001	0.14
Watched and listened to	-0.012	0.14

2.4.4 Correlation Analysis of Individual Differences (EDA_d Values).

Participants' EDA scores related to Study One were analysed for relationships with each of the individual differences (see Appendix L for descriptive statistics) under investigation to examine any meaningful patterns or trends, however none of the variables were found to be statistically significant using Pearson's correlation co-efficient analysis (see Appendix R).

2.4.5 Regression Analysis of Individual Differences (EDA_d Values).

The association of individual differences to EDA_d scores were explored via three multiple regressions, with the observed variable conditions; a) watched, b) listened to, and c) watched and listened to as outcome variables, and the nine individual differences measures (AEI experience [x2], AEI belief [x2], Unusual Experiences, Cognitive Disorganisation, Introvertive Anhedonia, Impulsive Nonconformity, and CPES) as predictor variables (see appendix L for descriptive statistics). All assumptions for multiple regression (linearity, non-multicollinearity, residual independence and normal distribution, homoscedasticity, absence of influential cases, and sufficient power) were met. Of the three multiple regressions, all returned non-significant models (see Appendix S).

2.5 Study One Results (Detection Measured by Task Performance)

2.5.1 Analysing the Results of the Task Performance Measure.

Whether participants' cognitive task performance would be affected by covert surveillance was tested by them taking an online Stroop Test (Ridley, 1935) under the four surveillance conditions (no surveillance, listened to, watched, watched and listened to). All participants completed the test without making any mistakes, and so only the response time was used in the analysis of Stroop Test performance. The descriptive statistics of the participants' task completion times (see Appendix T) are shown below.

Table 11

Task Performance Descriptive Statistics (Measured in Milliseconds)

Task performance conditions	Mean	Std. Deviation
No surveillance	1223.16	268.16
Listened to	1241.84	228.15
Watched	1296.52	289.98
Watched and listened to	1396.94	427.63

2.5.2 Analysing the Task Performance Scores Using a One-Way ANOVA.

A one-way repeated measures ANOVA was conducted to determine whether there was a statistically significant difference in task performance across conditions (no surveillance, watched, listened to, and watched and listened to). Mauchly's tests of sphericity indicated that the assumption of sphericity was violated, X^2 (5) = 51.65, p = .00004. Epsilon was 0.78 as calculated according to Greenhouse and Geiser (1959), and was used to correct the repeated measures ANOVA. There was a statistically significant difference in task performance across conditions F(2.36, 260.92) = 9.43, p = .00004, partial w^2 = .078 (see Appendix U).

2.5.3 Analysing the Task Performance Scores Using Paired-Samples T-Tests.

The test data passed most of the assumptions required (see Appendix V) for one-sample and paired samples t-tests, but the data was skewed. Due to the large sample size however, it was decided to still use t-tests rather than Wilcoxon rank sum analysis as with a sample of over 100 responses t-tests remain robust enough to be valid for any distribution (Lumley et al., 2002).

Paired-samples t-tests (see Appendix W) were conducted to investigate whether the delay in responding when there was an incongruent word-colour condition (for example if the word 'red' appeared on the screen but did not match the colour in which it was written) differed statistically depending on the surveillance condition (listened to, watched, or both) when compared to the control condition (no surveillance).

PAIRING 1: Listened to / no surveillance. A paired-samples t-test was conducted to compare how long participants took to complete the Stroop Test whilst they were under no surveillance, against how long it took them when they were being listened to. The difference in the time taken between these two conditions was not significant; $t(111) = -.72$, $p = 0.470$ (two-tailed), $d = .26$ (-0.01, 0.53).

PAIRING 2: Watched / no surveillance. A paired-samples t-test was conducted to compare how long participants took to complete the Stroop Test whilst they were under no surveillance, against how long it took them when they were being watched. The difference in the time taken between these two conditions was significant, with participants taking longer to complete the Stroop Test if they were being watched, compared to when they were not under surveillance; $t(111) = -2.15$, $p = 0.034$ (two-tailed), $d = .07$ (-0.18, 0.34).

PAIRING 3: Watched and listened to / no surveillance. A paired-samples t-test was conducted to compare how long participants took to complete the Stroop Test whilst they were under no surveillance, against how long it took them when they were being watched and listened to. The time taken for the participants to complete the task whilst being watched and listened to was significantly longer than when they were under no surveillance; $t(111) = -3.96$, $p = 0.001$ (two-tailed), $d = .48$ (0.22, 0.76).

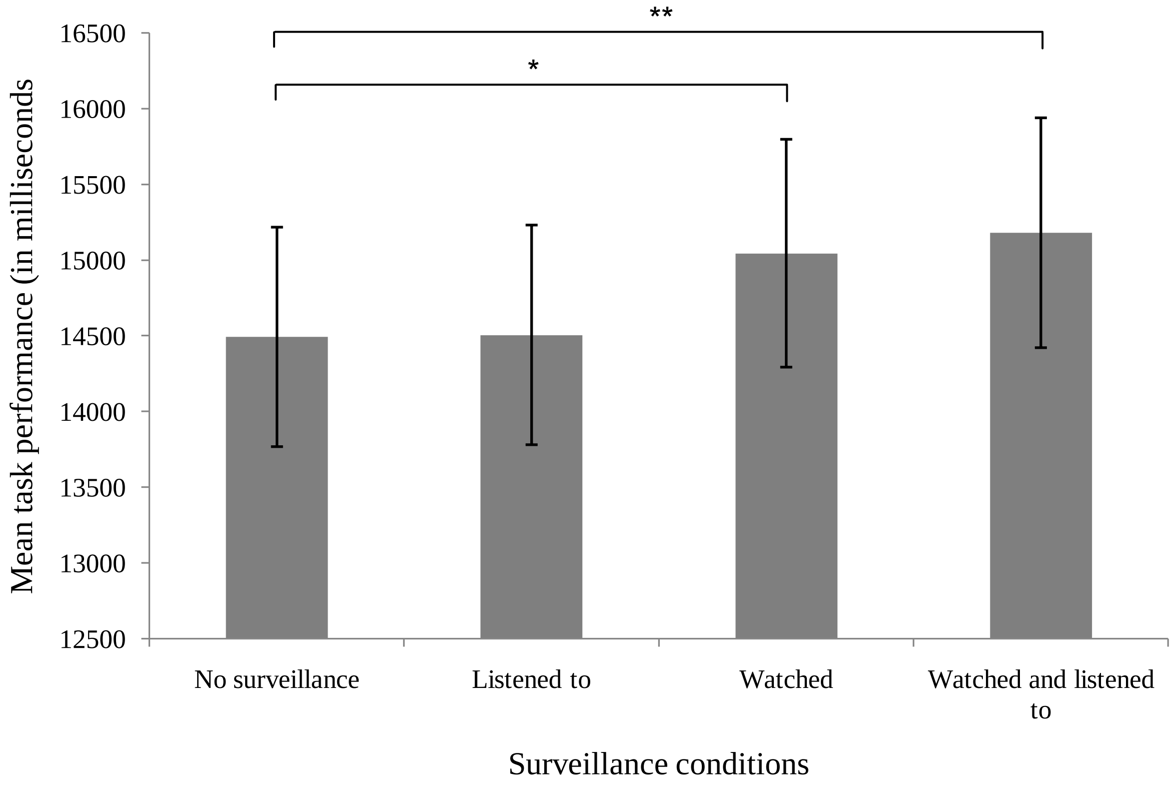


Figure 2 Significant task performance pairings.

Note. Paired-tests revealed that two of the surveillance conditions were significantly different to the control group.

* One-sample t-test significant at the 0.05 level (two-tailed).

** One-sample t-test significant at the 0.01 level (two-tailed).

2.5.4 Correlation Analysis of Individual Differences (Task Performance).

Participants' task performance scores related to Study One were analysed for relationships with each of the individual differences (see Appendix L for descriptive statistics) under investigation to examine any meaningful patterns or trends. The below variables were found to be statistically significant using Pearson's correlation co-efficient analysis (see Appendix X).

Table 12

Correlations Between Participants' Individual Differences Variables and Task Performance Scores

N = 112	No surveillance (*M* = 1223 *SD* = 268)	Listened to (*M* = 1241, *SD* = 228)	Watched (*M* = 1296, *SD* = 289)	Watched and listened to (*M* = 1396, *SD* = 427)
AEI Experience (*M* = 9.11, *SD* = 5.37)	-.241*	-.092	-.158	-.021
AEI Special Experience (*M* = 10.42, *SD* = 5.568)	-.230*	-.095	-.133	.006
AEI Belief (*M* = 5.02, *SD* = 2.92)	.016	.070	.028	.225*
AEI Special Belief (*M* = 6.43, *SD* = 3.368)	.003	.067	.042	.254**
Unusual Experiences (*M* = 4.49, *SD* = 2.71)	.039	.066	.026	.217*
Cognitive Disorganisation (*M* = 4.65, *SD* = 2.88)	.143	.182	.095	.199*
Introvertive Anhedonia (*M* = 2.95, *SD* = 2.42)	.052	.106	.107	.185
Impulsive Non-Conformity (*M* = 4.01, *SD* = 5.37)	-.172	.051	-.115	.071
CPES (*M* = 4.49, *SD* = 3.285)	-.047	.120	.044	.336***

*   Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

*** Still significant after the Bonferroni correction.

2.5.5 Regression Analysis of Individual Differences (Task Performance).

The association of individual differences to task performance score were explored via four multiple regressions (see Appendix Y), with the observed variable conditions; a) task performance while under no surveillance, b) task performance while listened to, c) task perfromance while watched, and d) task performance while while watched and listened to, as outcome variables, and the nine individual differences measures (AEI experience [x2], AEI belief [x2], Unusual Experiences, Cognitive Disorganisation, Introvertive Anhedonia, Impulsive Nonconformity, and CPES) as predictor variables. All assumptions for multiple regression (linearity, non-multicollinearity, residual independence and normal distribution, homoscedasticity, absence of influential cases, and sufficient power) were met. Of the four multiple regressions, only two returned non-significant models, the exceptions being task performance while watched and listened to, and task performance while under no surveillance.

For task performance while watched and listened to, the model (nine individual differences measures) explained 28.4% of the variance and was a significant predictor of task performance while watched and listened to $F(9,102) = 4.50$, $p = .000054$, with only the Introvertive Anhedonia ($B = .17$, $p = .05$) and CPES ($B = .60$, $p = .000381$) contributing significantly to the model.

For task performance while under no surveillance, the model (nine individual differences measures) explained 16.8% of the variance and was a significant predictor of task performance while under no surveillance $F(9,102) = 2.28$, $p = .022$, with only Impulsive Nonconformity ($B = -.20$, $p = .05$) contributing significantly to the model.

2.6 Discussion of Study One

It was hypothesised that participants' self-reporting of covert surveillance accuracy would be significantly greater than chance expectation and that correct responses would be significantly greater than incorrect responses, that EDA arousal levels will be significantly different when participants were under covert surveillance compared to when they were not, and that performance on the Stroop task will be different under surveillance conditions. The exploratory hypothesis that participants' ability to detect covert surveillance would be correlated with the individual differences variables was also tested. Analysis of the data revealed that some of these hypotheses were met, whilst others were not.

Analysing participants' ability to successfully self-report being watched and/or listened to was demonstrated to be no different to chance expectation regardless of the surveillance condition, and there was no significant difference between participants correct self-reports and their incorrect self-reports. Whilst Pearson's correlation analysis initially appeared to show that participants' Unusual Experiences, Introvertive Anhedonia and CPES scores had a relationship with self-reports - when a Bonferroni correction was applied to control for multiple analyses, none of these factors remained significant. Similarly, none of the multiple regressions performed returned significant models, and so none of the hypotheses related to self-reports were supported.

When testing the hypothesis related to whether participants EDA would prove a reliable indicator of surveillance detection, an ANOVA demonstrated there were no statistically significant differences in participants' mean EDA levels, no matter which of the surveillance conditions were analysed. As it could be considered meaningless to look at general magnitude with factoring in intra-individual differences from a no observation baseline, EDA_d were calculated and analysed via Pearson's for relationships with each of the individual differences, however none of the variables were found to have a statistically significant relationship. Multiple regressions were performed upon the EDA_d data, however none returned significant models. These analyses demonstrated that, as with the self-report measurements, none of the hypotheses related to EDA were supported.

With both the self-report and EDA measurements returning results that were not significant, the task performance data was examined using a one-way repeated measures ANOVA to determine whether the time participants took to complete their task correctly was affected by the surveillance condition. There was a statistically significant difference, and so paired-samples t-tests were subsequently conducted to compare the length of time participants took to complete their task under the various surveillance conditions, compared with the control condition when they were under no surveillance. It was found that participants took significantly longer to complete their task whilst they were being watched, and whilst they were being both watched and listened to, when compared with control - a finding that supports the hypothesis specifically related to the task performance measure. To investigate the exploratory hypothesis that participants' ability to detect covert surveillance would be correlated with the individual differences variables - a Pearson's correlation analysis was conducted. This revealed relationships between seven of the predictor variables under investigation, however only the CPES variable still showed a significant relationship after applying a Bonferroni correction meaning the exploratory hypothesis was only partially supported.

The association of individual differences to task performance scores were also explored via four multiple regressions, however only two returned non-significant models - namely task performance while watched and listened to, and task performance while under no surveillance. For task performance whilst watched and listened to, the model explained 28.4% of the variance with only the Introvertive Anhedonia and CPES contributing significantly to the model. For task performance whilst under no surveillance, the model explained 16.8% of the variance with only Impulsive Nonconformity contributing significantly to the model. Whilst this was an interesting finding, it again only partially supported the exploratory hypothesis.

Whilst the researcher remained prepared to accept that scopaesthesia may not exist, and that this may account for the hypothesis related to self-reports not being supported, alternative explanations were also considered. Previous research has suggested that even if genuine, the scopaesthesia effect is subtle and typically results in a success rate of around 55% (Sheldrake, 2003). Consequently, the well-established method of asking an individual to consciously guess whether they are being stared at may be valid, but the current self-report experiment suggests the signal is weak - if it exists at all. Indeed, an accurate staring detection rate of 54.5% was found, but this is in contrast with the 54.2% false positive rate. Interestingly, this result is in line with the work of Lobach and Bierman (2004) who found that many of their participants scored more positively than negatively, but not with statistical significance.

Of the three measures used in the current study as a means of detecting remote surveillance, the literature would arguably suggest that self-reports would be the most likely to produce positive results following over a century of research (Titchener, 1898) and the thousands of trials (Sheldrake, 2003) which appear to show that scopaesthesia is possible. Colwell et al. (2000), Radin (2004) and Sheldrake (1998, 1999, 2000, 2001, 2005b) have all argued that whilst they have found participants to be able to detect the presence of a stare, they are less able to detect its absence - this was shown to be the case in the current study, as whilst participants' detection scores were not statistically different to chance in the 'no surveillance' condition - they were in the predicted direction.

Prior to the experiment, the researcher had extrapolated from the body of work supporting scopaesthesia that if an individual may be able to detect being watched, they may also be able to detect being listened to. This was exploratory research as there are no studies on which to base this hypothesis, and the self-report data did nothing to strengthen the case for the idea of acoustathesia (Friday & Luke, 2014). Indeed, the results suggest that participants were not aware of when they were being listened to, and that they appear to have been guessing randomly when this was occurring.

Marks and Colwell (2000) have previously stressed the importance of eradicating the opportunity for sensory leakage and ensuring trials are properly randomised in research focusing on remote detection, and preliminary analysis of the current study's self-report scores appear to suggest that if this is considered and controlled for, the participants' ability to accurately self-report when they are being observed is rendered ineffective. However, to further investigate the results of the current study, the researcher also considered that the lack of positive results could possibly be accounted for due to a reporting bias - an issue repeatedly identified by Sheldrake (1998, 1999, 2000, 2001, 2005b). Analysis of the data showed that participants were indeed significantly more likely to answer 'yes' when asked if they could detect being watched or listened to, doing so on 53% of occasions. It could be argued that this over-reporting is likely to be related to expectation generated from the participants being told the aims of the experiment, and that surveillance would be taking place - and so the researcher considered that this possible barrier to self-reporting accuracy should be addressed in any subsequent related work or replications.

When considering whether to continue to pursue self-reporting as a means of covert surveillance detection, the experimenter noted that the data did indeed show trends in the predicted direction for all conditions - albeit not at a significant level. This is consistent with patterns suggested by Sheldrake (2003), and hints at a possible effect. Self-reports were therefore deemed worthy of further investigation. A plausible explanation for all conditions showing trends in the predicted direction, but without reaching statistical significance is that there could be a power issue, as when the data is combined across all conditions the effect size increases (and *p* value decreases), but again - not to a statistically significant level. With this in mind, it was determined that any further research in this area should include either more trials, or more participants.

Whilst considering further the limitations of the current study specifically with regards to self-reports, many similar studies that showed positive results had a greater number of participants such as those run at the NeMo Science Centre between 1995 and 2002 which saw more than 18,700 observer/participant pairs take part. Whilst the highly significantly positive results achieved during this experiment could be attributed to the lack of supervision and experimental control (Sheldrake, 2005), the outcome may have been due to the sample size which was approximately 167 times larger than the current study. It is worth noting however, that the participants' success rate was actually around 53%.

The experimenter also considered the apparent reporting bias a potentially important discovery. Whilst the current study found the accuracy of self-reporting to not be statically different from chance, the relationship between self-reports and accuracy was positive in every condition (no surveillance, listened to, watched, watched and listened to) and so further investigation is required as this seems potentially meaningful. It should be noted that answering positively in the 'no surveillance' condition does not support the case for remote surveillance detection as this is a false positive effect. Indeed, the various differences between the current study and others could account for the opposing outcomes - for instance Colwell et al. (2000) ran more trials, and Schachter (1959) used participants who had varying relationships with the observer, whereas Sheldrake (1998) did not vary the observer at all.

Whilst the results of the tightly controlled laboratory experiment related to the self-report protocol were at odds with the high levels of belief (between 68% and 97%) in remote observation among the general public (Braud et al., 1990; Braud et al., 1993a; Coover, 1913; Cottrell et al., 1996; Sheldrake, 2003; Thalbourne & Evans, 1992), it was posited that these results may perhaps contribute to an explanation of this difference. If the reporting bias was a contributing factor to the lack of significant results, the same could occur in everyday life when people believe they are being watched more often than they are, but tend to remember the times they were correct due to confirmation bias (Wason, 1960).

Some studies have found positive results when using specially selected observers, whilst the current study varied them in a concerted effort to avoid the observer effect which seemed to account for the powerful results in Schlitz and Wiseman's (1997) research. Follow up experiments could actively recruit for believers and supporters of the phenomena to investigate whether this would affect the accuracy of self-report scores, although the majority (71%) of the current study's participants reported experiencing staring detection, and 73% believed it is possible.

It has also been shown that practice and feedback regarding participants' accuracy has had a dramatic effect on participants' successful self-reporting (Colwell et al., 2000) as this allows them to confirm what the sensation of being stared at feels like to them. They can therefore supposedly recognise it when it occurs again. Similarly, Sheldrake (1998) found an accuracy rate of 90% when feedback was given to child participants (aged 8 to 9), so it could be suggested that practice and feedback may be a worthwhile consideration for future experiments, although it would also mean that any replications conducted as part of a second study would not be directly comparable with the first and so may be better deferred to a study designed to take this element into account from the start.

Despite there being possible explanations for the hypothesis related to self-reports not being met, there were no significant self-report results following Study One. As such, this element of the experiment would therefore be considered to support sceptics of scopaesthesia who have suggested that once the opportunity to use conventional senses to detect movement, sounds, shadows and reflections are removed, people's apparent ability to consciously detect when they are being stared at during a properly controlled and randomised experiment is diminished.

With the self-reporting aspect of Study One analysed, the researcher's attention turned to the second protocol which looked into whether participants' EDA may provide a valuable method of measuring remote surveillance detection. Whilst the self-report element of the current experiment yielded no significant results in Study One, researchers using psychophysiological measures such as EDA as a means of remote observation detection have found convincing evidence (Colwell et al., 2000; Peterson, 1978; Sheldrake, 2000; Thalbourne & Evans, 1992; Williams, 1983), whilst others have found their results to not be significant (Lobach & Bierman, 2004; Müller et al., 2009; Wiseman and Smith, 1994) in line with the current study's findings. Interestingly though, some researchers have even found the significance of their results to depend on which researcher ran the experiment (Wiseman & Schlitz, 1998).

Lobach and Bierman's (2004) experiment in particular provided some valuable insight into methods of measuring remote observation detection as it was the first of its kind to directly compare self-reports or 'conscious guessing' and EDA simultaneously. As the researchers found that EDA readings were a statistically significant 1% higher when under observation, it was predicted that the current study would find stronger results using EDA to measure surveillance detection than they would do using self-reports. In which direction the EDA values would change was more difficult to predict though as Braud et al., (1993a, 1993b) found that under certain conditions, observation can have a calming effect on participants' sympathetic nervous system and so a two-tailed analysis was used in the current research.

This physiological method of surveillance detection is not open to the reporting bias inherent with self-reports, and potentially serves as a way to capture unconscious and therefore objective detection of surveillance as the participant is not required to take action. This method is however prone to artefacts which can account for changes in the EDA as warned by Schmidt and Walach (2000). For example, the within subject variance in the current study had a standard deviation of 13-14%; this is a potential obstacle to detection if a seemingly genuine signal is a difference of only 1% (as reported by Lobach & Bierman, 2004). There are other potential issues which include the participant moving during the trial, or disturbances creating an arousal that may later be interpreted when analysing the data as a response to a change in surveillance condition.

The importance of staying as still as possible was explained to every participant taking part in the current study though, and distractions were kept to minimum with sessions being abandoned or re-started should this occur - despite all the care taken, a one-way repeated measures ANOVA demonstrated that none of the EDA means differed from one another with statistical significance. Whilst this could be due to a genuine lack of effect and there being no signal to detect, the results do not necessarily mean that measuring psychophysiological responses is a redundant way of detecting surveillance though, as in additional to issues concerning power - there are other feasible explanations such as the number or type of participants, the quantity or duration of trials, or the participants themselves - the signal may simply be too small to detect by the equipment used, or participants' EDA may not be sensitive enough.

It could also be speculated that there was a low signal to noise ratio due to artefacts, and too little power to detect such a signal. It was also considered noteworthy that the EDA results in the current study were 1.2% lower in the 'watched and listened to' condition when compared to the 'no surveillance' condition. Whilst this 1.2% was not a statistically significant difference, it was deemed a potentially valuable observation that was at least worthy of further investigation - largely due to the surveillance conditions involved. The researcher reasoned that if scopaesthesia (Sheldrake, 2003) exists, then the greatest difference in EDA means would be between the surveillance conditions that include being watched, and the condition during which the participants were not being observed. Furthermore, whilst theoretical due to the non-existence of research in relation to the idea of acoustathesia (Friday & Luke, 2014) - if this 'sense of being listened to' has any effect at all, it is reasonable to imagine that the condition in which participants were being watched *and* listened to in unison would elicit the greatest EDA response due to the doubling-up of attention focused toward the participants. It then logically follows that it would be the condition of being watched and listened to, compared with the condition under which no surveillance takes place that would result in a greater difference between participants' mean EDA scores than any other condition comparison. Therefore, because of the 1.2% difference found, it was considered that further exploration relating to this means of covert surveillance detection could be justified.

As the initial EDA means did not take individual variances into account, new EDA_d values were created to enable change within the participant relative to their 'no surveillance' EDA mean to be measured as it could be considered meaningless to look at general magnitude with factoring in intra-individual differences from a no observation baseline. These EDA_d figures were analysed for relationships with each of the individual differences under investigation to investigate any meaningful patterns or trends, however none of the variables were found to be statistically significant using Pearson's correlation co-efficient analysis. Similarly, the association of individual differences to EDA_d scores were also explored via multiple regressions, but all returned non-significant models.

Whilst previous experiments suggest that EDA provided a measurable method of detection, it could be that the situation itself may play a role in all of the detection measures as EDA is a direct response to stress (Silverthorn et al., 2009) - so perhaps the lack of threat in the current study rendered this measurement redundant in the absence of actual danger. This idea would make sense when considering the anecdotal evidence for scopaesthesia provided by Sheldrake (2003a) by Special Forces operatives, snipers, security guards and police officers who claim to be able to know when someone is looking at them and vice versa as these people who purported to use such abilities would need them for survival and protection.

The very fact that this was a laboratory-based experiment and time and attention had been taken to make the participants feel welcome, safe and secure may be responsible and ethically necessary - but could also be an obstacle to such research as this may act like an 'off' button for extrasensory ability, supposedly developed to warn of imminent danger (Sheldrake, 2003). This idea is supported in the literature by Robbins and Parlavecchio (2009) who used the Unwanted Exposure Model of Embarrassment to explain that embarrassment is elicited when the core self has been exposed to unwanted attention. Therefore, it is reasonable to suggest that the 'unwanted' attention created by the remote staring effect - particularly if the participant was unable to 'interconnect' with participants who would be remote observers (Braud & Shafer, 1993), would elicit feelings of embarrassment when the threat of being watched risks exposing the participants' core self to an unfamiliar stimulus.

The idea that threat, danger or embarrassment could play a role in automatic biological responses to being watched as a response to threatening stimuli or danger is also supported by Gerlach et al., (2003) who showed that exposing participants to a watchful audience produced high rates of self-reported feelings of embarrassment, as well as increased sympathetic activation. Therefore, not only did participants report feelings of embarrassment to being watched, participants also showed automatic biological responses to being watched, which indicates that feelings of embarrassment may be an evolutionary response to the threat related stimulus of being stared at (Gerlach at al., 2003).

Another important element to consider is that familiarity may play a role in moderating feelings of embarrassment as a response to danger. Braud and Shafer (1993) identified that by allowing participants to become familiar with one another before the remote staring experiment enabled participants to feel positively towards being watched. Therefore, feelings of embarrassment may only be elicited during remote staring when the observer and the observed are unfamiliar to one another (Braud & Shafer, 1993). Importantly, in Study One the participants met at the start of the experiment. A study by Dahl et al., (2001) supports the role of familiarity in moderating embarrassment and threat after finding that consumers who were familiar with an embarrassing product were less embarrassed to purchase the product due to the high level of familiarity between the consumer and the product. These studies not only serve as potential explanations for the lack of significant results relating to the EDA element of Study One, but also to suggest how future methodology could be evolved.

Whilst it would of course be impossible to re-create a life-threatening or dangerous situation to investigate whether this is required for individuals to 'engage' their ability to detect surveillance, an ethically acceptable alternative may be possible and could be considered for future experiments. Indeed, the literature suggests that embarrassment, or the involvement of strangers could provide ethical, but viable options.

It is also possible that the concentration of the observer participant watching and listening to their counterpart could have played a role as suggested by Schwartz and Russek (1999) who noted that observers' intention or attention may be essential part of their stare being detectable. Whilst the randomisation of condition sequences employed in the current experiment would have ensured that no particular condition would have been affected by participant fatigue, the lack of a genuine need to attend fully to the task at hand may account for the apparent lack of detection. Beyond altruistism or a sense of duty - neither the observer nor the observed participants had any real incentive to focus and dedicate their attention to the target, and so their concentration may have waned during the surveillance time, if indeed it was ever gained. Similarly, whilst instructed to concentrate on whether they were under surveillance, the participant whose EDA was being monitored could also have not done so as they were required to do nothing other than sit still, and so their dedication to the aims of the experiment would be impossible to access. A lack of attention to whether they were under surveillance or not may account for the lack of significant results.

Another possible explanation for the lack of EDA response could also lie in the literature as EDA seems particularly susceptible to the experimenter effect as evidenced by the work of Wiseman and Schlitz (1998) when significant results seemed to depend entirely on which experimenter conducted the studies, as the experiments were otherwise identical. Wiseman is a renowned sceptic and Schlitz is a believer in extrasensory abilities, and both researchers found results in the direction they would expect. This experimenter effect also seemed to be evident in relation to measuring surveillance in experiments conducted by Braud et al. (1993a, 1993b) and Wiseman and Schlitz (1999), and so it seems the observer themselves can be a factor. As the current research recruited participants randomly, the sample contained both believers and non-believers. Perhaps a follow up experiment that recruited specifically for participants who believe in such abilities may provide significant results.

Whilst it was hypothesised that the current experiment would find that surveillance conditions would have a significant effect on participants EDA due to Braud et al. (1990; 1993a) concluding via their carefully controlled experiments that EDA may be a better measure of detecting remote observation than conscious guess measurements, as these authors suggested that unconscious measures may be more sensitive, and so may produce a better means of measuring remote observation (see section 1.3.4 for further details). It may be important to note that their study was based on sixteen participants, and that perhaps this effect may have disappeared had they used as many participants as the current study. Their research was supported by similar studies conducted by Schlitz and LaBerge, (1994, 1997) which had a greater number of participants, but these researchers had their participants observed using CCTV, which is not the case with the current study meaning it is not necessarily directly comparable.

This difference in observation methods is potentially an important one as Schlitz and Braud's (1997) meta-analytic review of 11 studies encompassing 230 individual sessions which used psychophysiological means to detect remote observation found that 65% demonstrated a significant difference in remote observation detection between observation and non-observation trials - and all were via CCTV, and so electronic surveillance may be a direction for future experiments. It should be considered however that if CCTV elicits the ability to detect remote surveillance whilst direct surveillance does not, this could be argued to oppose the theory of scopaesthesia's origins. The phenomenon is proposed to be evolutionary and to date back to when humans were both hunters and the hunted, whilst CCTV is a comparatively recent invention. Conversely, a counter argument is that this finding actually *supports* evolutionary theory (Sheldrake, 2003) as it highlights the possible mode of mechanism - namely that attention and intention is more essential to scopaesthesia than line of actual sight as suggested by Sheldrake's 'two-way process' theory (1994).

Less studied in relation to surveillance detection than either self-reports or EDA is task performance. The idea to include it in the current study was based on the social psychology idea of reactivity (Heppner et al., 2008) which occurs when individuals alter their behaviour due to an awareness that they are being observed. This is similar to the well-researched area of social facilitation which is described as peoples' tendency to behave and perform differently when alone to when they are in the presence of others (Griffin & Kent, 1998). Using this idea, the researcher hypothesised that if people perform tasks differently when under conventional observation, then they may do so when the observation is covert if they are detecting its presence via unconventional means. They also originally hypothesised that if remote surveillance influences psychophysiology, then it is likely to influence task performance - however when the current study found that both the self-report and EDA measurements suggest that none of the surveillance conditions are significantly detectable by participants, theoretically task performance became a less likely means of detection. This is because social facilitation dictates that individuals respond to the knowledge they are being observed by altering their performance and behaviour, therefore if the self-report measure has shown them to be unaware of surveillance, and the EDA measure demonstrated that they are not unconsciously aware either - it seems illogical that their performance could be affected by surveillance they could not be aware of via conventional senses.

Whilst changes in performance due to remote surveillance have not previously been explored in combination with the other methods of detection employed in the current study, the idea that remote observation may affect behaviour was investigated by Sheldrake (2003). He examined how many people would turn to face a CCTV camera when it was monitored, as opposed to when it was unmanned and found individuals did so 27 times during the observation periods, and just 12 times during the non-observation periods. This is a highly significant difference in the reactions of the people under observation, and whilst this is the only known study to find positive results when examining this effect - it does suggest that behavioural responses to surveillance should at least be considered.

Surprisingly, whilst due to the limited research in this area, the inclusion of a task performance measure was to some extent exploratory as it has been less researched than self-report or physiological measures - it provided the only statistically significant results of Study One. A one-way repeated measures ANOVA was conducted to determine whether a significant difference in task performance existed, and demonstrated that this was indeed the case. Subsequent paired-samples t-tests were used to analyse whether participants took significantly longer to successfully complete the Stroop Test under the various surveillance conditions compared with the control condition (no surveillance). Whilst there was no significant difference in the total delay in decision making measured by milliseconds during the Stroop Test when comparing participants' under no surveillance with when they were listened to - participants took significantly longer to complete the Stoop test under the 'watched' and the 'watched and listened to' conditions, than when they undertook the same test whilst under no surveillance at all. These findings favour the genuine existence of scopothesia relative to acoustathesia when task performance is considered.

This preliminary analysis of task performance as an indicator of covert surveillance detection suggests that remote attention from others has a negative effect on individuals' task performance capability. Task performance could therefore be considered a possible measurement of this phenomenon, and is not only potentially evidence of scopaesthesia, but it could also be argued to support the idea of acoustathesia (Friday & Luke, 2014) as it would appear that when the participant was listened to as well as watched, the size of the effect was greater than when they were just watched. It should be noted though, that no test of significance between these two conditions was performed to reduce multiple analyses.

The bias which seemed to have affected the results of the self-reports could not have played a role with this task performance measurement as participants were not choosing a response, and the results cannot be attributed to practice effects or participant fatigue as the condition sequence was randomised for each participant. The effect may have been exaggerated though had the participants not been introduced to each other and made to feel so welcome and relaxed at the beginning of the experiment as a lack of familiarity between those involved has been shown to alter the observed participants' behaviour (Guerin, 1993) with conventional reactivity.

Braud and Shafer (1993) found that promoting 'inter-connectedness' between participants who would come to familiarise themselves with one another related to a participant's ability to enjoy the experience of remote staring, and they were then able to experience the sensation of being watched while feeling relaxed. Conversely, participants who were not able to connect with fellow participants before taking part in the remote staring experiment reported that the sensation of being watched was more threatening. For individuals who were unfamiliar with fellow participants, the experience of being stared at from an unseen person heightened the perceived sensation of danger (Braud & Shafer, 1993). Despite these emotional differences in remote staring conditions from individuals who were either interconnected or not interconnected with fellow participants, the same level of accuracy in detecting remote staring at a sub-conscious level was found by both groups.

In a replication of Braud and Shafer’s 1993 study, the role of personality was explored in heightening the remote staring effect. Individuals who displayed high scores in the Social Avoidance and Distress scale (SAD) which measures social distancing, distress and anxiety - and who therefore may lack interconnectedness/familiarisation with other participants, showed greater sensitivity to detecting staring from an unseen individual. This indicates that perhaps a lack of familiarity with an individual, due to prolonged isolation which relates to SAD personality types, may elevate an individual’s vulnerability to the remote staring effect (Braud et al., 1993).

Braud and Shafer (1993) argued that a person who exhibits high scores in SAD may show an increased desire to interconnect with other individuals, which may in turn make individuals who are high in social distancing/distress/anxiety more sensitive to remote staring detection, as they try to increase their connectedness to other people that they may not yet be familiar with - even in non-remote staring situations. According to Bond and Titus (1983), if the pair of individuals are familiar with one another, the social facilitation effect will be reduced. Therefore, the fact that the majority of the participants were strangers in the case of the current study may have contributed to the significant results. Alternatively, the welcoming atmosphere and introductions at the beginning of the experiment may have reduced participants’ sensitivity to detecting covert surveillance. The researcher therefore considered that the positive effects may have been exaggerated further if the participants were kept completely separate at the start of the experiment, with no opportunity to familiarise themselves with each other. This idea was considered for future research and partial replications.

A further contributing factor to the significant results achieved via the task performance measurement may have been the degree of attention the participant watching and listening to their counterpart dedicated towards them. Horrey et al., (2017) used the example of a driving task to demonstrate how attention is greater for an interesting task as opposed to an uninteresting task. During the driving task, participants reported that listening to interesting material while driving was less demanding and less complex than uninteresting material, despite both materials being matched for difficulty. Participants also demonstrated reduced concentration of cerebral oxygenated haemoglobin when listening to interesting material compared to uninteresting material, and still managed to show superior recognition for the interesting material (Horrey et al., 2017). Therefore, an interesting task may generate greater attention from an individual as the task itself may be less trying or demanding, and, from a psychobiological perspective, may reduce pressure for the individual to concentrate due to the relative ease of an interesting task (Horrey et al., 2017).

The lack of attention given to a mundane or boring task is further supported by the psychobiological perspective, where Danckert and Merrifield (2016) observed that boredom relates to an individual's inability to engage executive control networks when faced with an uninteresting task. Therefore, despite being presented with tasks that do require engagement, individuals may show a failure to do so when the task itself is mundane or boring (Danckert & Merrifield, 2016). The role of executive control networks in enabling an individual to focus on mundane tasks suggests that boredom may impact neurobiological factors that would enable individuals to pay attention to tasks, which highlights how the absence of boredom when performing an interesting task ensures greater attention and focus on certain tasks. The idea that the degree of attention the participant paid to the task at hand may have played a part in participants' ability to detect surveillance has been mentioned earlier in this discussion as a possible explanation for the lack of positive results, particularly in relation to the EDA element of the current study. Whilst tenuous, the author cautiously suggests the reverse may be the case here, and that attention may be a possible reason why surveillance appears to have been detected in the task performance study.

Whilst task performance itself may simply be a better detector of surveillance, this element of the experiment may have attracted greater concentration from the participant watching and listening. Not only is task performance the only measurement that requires constant activity from the participant under observation, therefore providing a stimulus to the observer which may hold their attention - but the total duration of the trials is shorter, with breaks in-between sessions to reset the task. This means that sessions of watching and/or listening are in periods of around just one minute. The observer participant is told whether they should watch, listen, or do neither at the start of each trial and so is not partially distracted by anticipating the next set of instructions as was the case with the self-report and EDA protocols. Whilst participants in all of the pilot studies reported that they found following the condition sequence instructions for the self-report and EDA measures simple, perhaps this was more distracting than they consciously realised.

With significant results found, the time it took participants' to complete the task was analysed for relationships with each of the individual differences under investigation to examine any meaningful patterns or trends. Whilst seven of the individual differences initially appeared to be related to the task performance scores (see section 2.5.4), only the CPES variable related to temporal lobe lability remained statistically significant after a Bonferroni correction was applied to take multiple analyses into consideration. Interestingly though, this was specifically the case in the 'watched and listened to' condition which the current study would suggest is the condition in which surveillance detection appears to be most accessible.

The association of individual differences to task performance scores were explored via multiple regressions, two of which returned significant models - these were task performance while watched and listened to, and task performance while under no surveillance. Again, the experimenter thought this may be meaningful as these two conditions are at either ends of the extreme (the least surveillance possible, and the most surveillance possible).

For task performance while under no surveillance, the model explained 16.8% of the variance and was a significant predictor of task performance, with only Impulsive Nonconformity contributing significantly to the model. Whilst the author concedes that the following is somewhat a logical leap, this thinking is supported by their earlier work (Friday & Luke, 2014) in which Impulsive Nonconformity demonstrated a similar effect - namely that the participants who scored highly on this negative trait appeared to be reporting *against* expectations. As such, the author tentively suggests that participants scoring highly on the Impulsive Nonconformity scale may have been deliberately reacting in the opposite direction to which they were expected to. However, the author is also keen to stress their awareness that if this were the case, it could be argued that the 'watched and listened to' and 'watched' conditions would have generated negative p values as the participants scoring highly on Impulsive Nonconformity would theoretically report not being under surveillance when they believed they actually were.

For task performance while watched and listened to, the model explained 28.4% of the variance and was a significant predictor of task performance, with only the CPES and Introvertive Anhedonia predictor variables contributing significantly to the model. The CPES scores sharing a relationship with task performance scores in the 'watched and listened to' condition in hint that temporal lobe lability may be a neurological predictor of surveillance detection as previous research would suggest (Horowitz & Adams, 1970; Luke et al., 2013; Persinger et al., 2000), however with this only being found in only one surveillance condition, within just one of the experiment protocols, the experimenter considered more evidence was needed before making such a claim.

The relationship shared with Introvertive Anhedonia was perhaps less expected, as this O-LIFE subscale is characterized by schizoid solitariness and a lack of feeling (Claridge & Beech, 1995), and has been interpreted as being characterised by a lack of enjoyment as a result of physical pleasure and social interaction, coupled with a withdrawal from both physical and emotional intimacy which is then replaced with an importance placed on solitude and independence (Mason et al., 1995). It could be argued that this latter characteristic would make high Introvertive Anhedonia scorers particularly sensitive to the being watched and listened to, as they may associate this with a sense of intrusion more than participants who scored lower on this scale. Again, more data would be required to support such a statement, and the problems with multiple analyses would need to be considered.

In summary, returning to the point made earlier that if surveillance is genuinely not being detected consciously as the self-report element in Study One seems to suggest, or unconsciously as the EDA protocol appears to indicate - then the apparent effect found via task performance is surprising. The author therefore cautiously suggests that perhaps surveillance may have been potentially detectable via *all* of the measurements investigated in Study One, but that the instruments used were either not sensitive enough, or that the methodology should be refined. As self-reports did appear to show a weak signal in the predicted direction, this measurement of detection could be pursued further - but with more trials, more participants, or varied periods of surveillance. These ideas could also be applied to the EDA measure, or an alternative method of detecting psychophysiological changes could be employed.

It is hoped that the tight controls that were fundamental in this study will contribute to the literature by addressing the problem highlighted by Schmidt et al. (2004) regarding possible artefacts with large datasets, and that it will inform the next experimental stage of the current research as well as the work carried out by other experimenters to determine whether there are correlations across measurements, and to investigate whether there is a detection tendency independent of modality. The results thus far appear to suggest that whilst being listened to may not as a condition in and of itself elicit accurate detection by the participants, the aforementioned ‘doubling-up’ of being watched and listened to in combination may be the most detectable condition. Should further evidence for this exploratory theory be discovered - it would be an original and important addition to the literature, and so the necessity for replication, albeit with methodological improvements seems clear.

3 STUDY TWO

3.1 Study Two Introduction.

Having conducted Study One and analysed the results, a partial replication with some methodological developments was deemed necessary for Study Two. The results from the first study were largely unexpected, and therefore led to new avenues of investigation for all three methods of surveillance detection tested (self-reports, EDA, task performance). So as to not re-design the initial study drastically and render the second study incomparable to the first, small but potentially important changes were introduced.

Whilst not at a significant level, the self-report results from Study One (see section 2.3.3) suggested that there may have been a weak signal guiding participants' 'guesses' as to when they were under surveillance, as whilst their accuracy rate was not statistically significant - the results were in the predicted direction in all conditions. This was therefore deemed worthy of further investigation as previous research has suggested that the scopaesthesia effect, even if genuine is subtle with a typical success rate of approximately 55% (Sheldrake, 2003). The literature on self-reports or 'conscious guessing' as a means of detection implies that even if the ability to detect surveillance exists, the signal is weak. Indeed, Study One's self-report experiment resulted in an accurate staring detection rate of 54.5%, in contrast with the 54.2% false positive rate. Whilst neither of these percentages are statistically significant, they support the research of Lobach and Bierman (2004) as these experimenters' results demonstrated that their participants also scored more positively than negatively, but were also not at statistically significant levels.

To establish whether the similarity between the results of the current study and Lobach and Bierman's (2004) research is meaningful, further experimentation was planned. It was concluded that the aforementioned trend of the self-report results being in the predicted direction whilst not reaching statistical significance could suggest a power issue, as if data for all conditions is combined, the effect size increases whilst the p-value decreases.

To investigate whether power may have indeed been an issue, the experimenter ran a power analysis. The power analysis for alpha = 0.5 and power of 80% suggested that an increase in participants from 112 to 313 individuals would be needed. Following the lengthy and difficult process of recruiting the participants for Study One, an increase to 313 participants was deemed unrealistic with the time and resources available. After careful consideration, the experimenter planned to recruit a similar number of participants for Study Two, but having them complete double the number of trials which would fall close to the 70% power bracket of N= 246. Whilst not a perfect solution, due to the limitations of the recruitment period and the participant incentives on offer - it was decided that this was the best available option.

To attempt to establish whether a lack of statistical power may account for the results of Study One's self-report protocol not concurring with the majority of the previous research, Study Two's self-report experiment used the same methodology, but with twice as many trials.

To try and tackle the issue of reporting bias, the experimenter considered various ways of running the experiment without informing the participants that periods of surveillance will occur - however ultimately it was deemed impossible for those taking part to consciously guess whether they are being watched and/or listened to, without their knowledge that this will take place as the computerised programme was asking them to guess. The experimenter did decide however to not highlight this by omitting to the participants that surveillance was part of the study until after the experiments concluded.

The current researcher's decision to include acoustathesia (Friday & Luke, 2014) which is described as the sense of being listened to was exploratory due to the lack of research on which to base this hypothesis. Whilst the self-report protocol from Study One did little to support this idea as the data appeared to suggest that participants were merely guessing randomly when this condition occurred, due to the positive findings related to the task performance measure (see section 2.5.3), the experimenter decided to pursue this surveillance condition further. The idea of acoustathesia was simply extrapolated from the body of work on scopaesthesia, but the results of the task performance protocol from Study One does seem to hint that the 'doubling-up' of being watched and listened to strengthens the scopaesthesia effect, as this combination appeared to be more detectable than simply being watched.

It was predicted that during the self-report protocol, the observed participant may deduce that surveillance could be taking place due to the necessity for them to actively report whether or not they felt like they were being watched or listened to. For this reason, rather than the three protocols (self-reports, EDA, task performance) being randomised as they were in Study One, throughout this partial replication they were deliberately ordered so that the self-report protocol was always the last to take place to ensure that *if* the unavoidable accompanying instructions gave away or inadvertently hinted to the participants that periods of surveillance would be taking place - any resulting expectation would not affect the EDA and task performance elements of the study. It was hoped that this methodological change would help eradicate, or at least lessen the reporting bias that affected the first study.

The EDA experiment in Study Two was also designed to address possible reasons for the lack of significant results (see section 2.4) related to physiology measures as a means of surveillance detection (in any direction, positive or negative) as the researcher initially hypothesised this unconscious response would accurately demonstrate participants awareness of being watched and/or listened to. Whilst the EDA results from the first study were 1.2% lower in the 'watched and listened to' condition than in the 'no surveillance' condition, the difference was not statistically significant. Whilst this could again be related to statistical power, or artefacts (a problem highlighted by Schmidt and Walach, 2000), none of the EDA means statistically differed from one another, regardless of the condition. To consider the possibility that the reporting bias which marred the self-report section of Study One may have masked what was already expected to be a subtle signal - the experimenter decided to replicate the initial EDA section of Study One again, but *before* advising the participant that they would be under surveillance in an attempt to eradicate the possible artefacts of guessing or predicting surveillance periods.

Whilst the experimenter acknowledges that the lack of significant results associated with EDA in Study One could be explained by the non-existence of a genuine effect, for the EDA protocol in Study Two, it was hypothesised that the engagement created by the Stroop test utilised in the task performance protocol may have played a role in participants' ability to detect surveillance. This idea is supported by Study One's task performance protocol results (see section 2.5.3) in which the participants being 'scored' may have encouraged them to care about their performance more, and so could be argued to have helped generate positive results. Study Two's EDA protocol was therefore designed to measure the participants' EDA responses whilst they undertook the same task that generated the significant results in Study One's EDA protocol. The notion that this change in situation could make an important difference is supported by research which states EDA is a direct response to stress (Silverthorn et al., 2009). However, Shafer (Braud et al., 1993a, 1993b) and Schlitz (Wiseman & Schlitz, 1999) demonstrated how remote staring may achieve a powerful physiological reaction in participants which is measurable by

EDA, but also showed how this could be either a calming or an activating influence - and so, as in Study One, no direction was predicted for the EDA results related to Study Two.

Indeed, the evolutionary explanation of scopaesthesia (Sheldrake, 2003a) also fits with this experimental re-design as the ability to detect being watched is theorised to emerge under intense situations such as hunting or being hunted. It is therefore worth noting that mood manipulation as part of emotional Stroop tasks has been successful at creating changes in self-reported state anxiety in participants (Richards et al., 1992). Individuals who exhibited high trait anxiety showed interference effects that were consistent with the induced mood produced by the Stroop task (Richards et al., 1992). Richards et al. (2000) also demonstrated that as anxiety increases, there was a similar increase in interference produced by threat related stimuli, thereby demonstrating that anxiety and interference in a Stroop task share a significant linear relationship.

When examining changes to heart rate and skin conductance responses, the Stroop task was found to be an efficient laboratory stressor (Renaud & Blondin, 1997), as participants' heart rates and ratings of state anxiety remained high throughout the Stroop task. However, Renaud and Blondin (1997) were unable to show a relationship between the induced mood created by the Stroop task and changes to skin conductance responses. Despite this, more recent studies have demonstrated a relationship between changes in EDA responses and the Stroop effect. For example, by using a virtual reality Stroop task (VRST), Parsons and Courtney (2018) were able to demonstrate that the Stroop effect related to increased skin conductance responses, elevated heart rate, and increases in spontaneous fluctuations in EDA. Parsons and Courtney (2018) highlighted a direct relationship between automatic changes in physiological arousal and VRST, which suggests that the threat-related stimuli evoked by the Stroop task helps to elicit feelings of anxiety that may lead to fluctuations in EDA, due to the automatic physiological changes the task produces. It is therefore interesting that Sheldrake's (2003a) anecdotal supporting evidence offered by Special Forces operatives, snipers, security guards and police officers who claim to use such extrasensory abilities would also need them for survival and protection. With creating real and actual danger within a laboratory-based experiment not a feasible possibility, it was considered that the participants being tested and their perception of being 'judged' on their scores might create an ethically acceptable alternative whilst creating the stress necessary for measurable EDA responses to surface.

Study Two provided a chance to investigate possible reasons why certain measures failed to demonstrate a significant means of surveillance detection in Study One, but it will also offered a chance to try and explore positive results further such as those gained by investigating task performance during Study One. These results appeared to suggest that differences in the time taken for participants to accurately complete a Stroop test (Ridley, 1935) is a demonstrable indicator of whether they are aware of the attention directed at them.

A simple recreation of the task performance protocol would be interesting in and of itself to investigate whether the initial results are replicable, but whilst the bias which appeared to have affected the results of the self-reports section should not have influenced the performance as the participants were not consciously choosing a response - running this part of the experiment again, but in combination with the EDA element could offer further insight. And by conducting this section *before* it is revealed to participants that periods of surveillance will be taking place, the experimenter will be able to directly compare the results of when the participants were, and were not aware of being ‘watched’ and/or ‘listened to’. If an identical replication was taking place, based on the results of Study One, a one-tailed hypothesis would replace the previous two-tailed hypothesis as the direction of change could be predicted. In Study Two however, the participants would not be told that surveillance would take place and the effect of this change is unknown, and so the hypothesis remained two-tailed.

3.1.1 Study Two Hypotheses.

Having evaluated the results of Study One, further hypotheses were formulated for testing in Study Two. These are detailed below.

Formal Hypotheses.

- Self-reporting of covert surveillance accuracy will be significantly greater for correct self-reports when compared with incorrect self-reports when the number of trials per participant is doubled, and priming is avoided.

- EDA arousal levels will be significantly different in the epochs when participants were under covert surveillance compared to when they were not if they are undertaking the Stroop test.

- The time taken for participants' to accurately complete the Stroop task will be greater under covert surveillance, compared to when not under covert surveillance.

Exploratory Hypothesis.

- The ability to detect covert surveillance will be predicted by belief in and experience of anomalous phenomena, temporal lobe lability, and Schizotypy scores, although this is an exploratory hypothesis investigating possible relationships between these factors when the participants are not primed to expect covert surveillance.

3.2 Methods

3.2.1 Design.

Study Two employed a multivariate mixed experimental design, and continued to examine which psychosocial and neurological factors may predict individuals' ability to detect being watched and/or listened to via extrasensory means. This was again conducted using four conditions i) no surveillance, ii) being listened to, iii) being watched, and iv) being watched and listened to. As in Study One, the variables of interest (belief in and experience of anomalous phenomena, temporal lobe lability, and levels of schizotypy) were measured via questionnaire before the experiment began. Participants' ability to detect surveillance was measured again using;

- Self-reports - although with the number of trials doubled from eight to sixteen per participant to increase power as in Study One the results were all in the predicted direction, but not at a statistically significant level.

- Changes in EDA whilst under (or not under) surveillance - although this time an additional protocol was conducted to measure participants' EDA whilst they undertook the same Stroop test utilised in the previous task performance protocol to investigate whether the experimenter's idea that the focus, attention, or threat it created may account for the significant task performance results.

- Changes in task completion time to add support to the findings in Study One that showed participants took significantly longer to complete the task under certain surviallance conditions.

All of the detection measurements were again taken during randomised periods of surveillance that the participants could not be aware of without extrasensory ability due to tight experimental controls. As before, participants' accuracy in detecting surveillance was analysed for possible relationships, trends and patterns with the individual differences variables under investigation - however, a potentially important difference was that in Study Two, the participants were not introduced to each other at the start of the experiment and were in fact completely unaware of one another's involvement in the experiment. This was because of a change in methodology to ensure participants were not aware surveillance would be taking place, intended to investigate whether the reporting bias which may have inadvertently been created by them knowing this information in advance during Study One played a role in their surveillance detection results. Following this change to the procedure and experimental design, it was expected that each measure would predict participants' ability to detect being watched and/or listened to via unconventional means. One-way ANOVA's, one-sample t-tests, paired-samples t-tests, Pearson's correlation analysis, and multiple regressions were again used to examine the resulting data for possible relationships, predictors and patterns.

3.2.2 Participants.

The participants (all aged 18+) were recruited via advertisements in the media (press and radio) and SONA (the universities participant recruitment system) which promoted the study to students at a London University. This time the materials did not mention that experiments concerned surveillance detection. This produced 121 respondents, of whom 110 proceeded to take part in the study. Due to sample size ($N > 50 + 8m$) requirements for sufficient statistical power (Tabachnick & Fidell, 2007), a minimum of 106 participants (see Appendix Z) were required for meaningful analysis. Participants consisted of 57 (52%) females and 53 (48%) males, ranging in age from 18 to 67 years ($M = 37.35$, $SD = 12.49$).

3.2.3 Materials.

The study's data collection was conducted in the same observation laboratory as Study One so as to keep the environment identical, and so that physical differences in the participant's surroundings could not account for any differences in results. For this reason, the apparatus and programmes also remained the same (for a detailed description see section 2.2.3).

The psychosocial and neurological factors under investigation were the same as in Study One. For full details and justification see sections 1.6.2, 1.6.3, and 1.6.4, and for a more detailed version of the following see section 2.2.3, however the questionnaires are briefly detailed again below.

Anomalous Experience and Belief. This measurement was again operationalized by the experience and belief subscales of Gallagher et al.'s (1994) Anomalous Experiences Inventory (AEI) scores (see Appendix C).

Temporal Lobe Lability. Temporal lobe lability was measured by Persinger and Makarec's (1993) Complex Partial Epileptic Signs (CPES) and Temporal Lobe Symptoms (TLS) subscales of the Personal Philosophy Inventory scores (see Appendix D).

Schizotypy. Schizotypy was measured by Mason et al.'s (2005) Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE) scores (see Appendix E).

As before, demographic questions recorded participants' age and gender.

Measures used to determine whether surveillance has been detected are listed below. However, as they are the same as in Study One and were therefore fully explained in section 2.2.3, so are simply listed here.

Conscious self-reports recorded by PsycLab.

Physiological reactions determined by EDA readings.

Behavioural change measured by the participants' performance whilst undertaking the Stroop Test (Ridley, 1935).

3.2.4 Pilot Studies and Experimental Amendments.

Whilst Study One presented challenges for the researcher, particularly with balancing the needs of two participants performing opposing tasks at the same time - the study did run efficiently following three pilot tests. For this reason, and to ensure that possible differences between the results of Study One and Study Two were due to intended alterations detailed in section 3.2.1 rather than procedural variations - Study Two was identical except for these three important changes. A pilot test was run for each to determine possible problems, and to identify the solutions required.

3.2.5 Pilot Study One (Doubling Self-Report Trials).

The amount of trials had been carefully considered for Study One's self-report protocol, however concerns of participant fatigue during an hour-long study which was already demanding in terms of attention and concentration led to the researcher deciding on just eight self-report trials per participant (surveillance twice, listened to twice, watched twice, and watched and listened to twice) in a randomised order which lasted 4 minutes in total (8 trials lasting 30 seconds each). Whilst increasing the number of trials to 16 would double the statistical power of the self-report results, it was important that this did not result in a decrease in participants' dedication to the task as their waning concentration may affect results. A pilot study was therefore conducted to establish whether eight minutes performing a mundane task would be detrimental to the study. The six participants did not report feelings of boredom or frustration, and their task performance score means for the second half of the trial were generally consistent with the first half, indicating that fatigue would not be an issue that would affect the validity of the results.

3.2.6 Pilot Study Two (Avoiding Reporting Bias).

In the self-report protocol related to Study One, the results were complicated by a reporting bias suspected to be caused at least in part by telling the participants that surveillance would be taking place - but the decision to delay revealing this information until after the experiments had finished came with methodological complications. In the first version of the self-report protocol, the order and manner in which the participants arrived was not of concern as they were literally introduced to one another. It would then be explained that one participant would be observing the other and vice versa, and so any movement of the experimenter between rooms to communicate or assist either participant did not need to be considered. This was however not the case for the second version of this experiment. Should such activities be noticed, glimpsed, overheard, or even suspected by the observed participant - the methodological change was effectively redundant, and the experiment was again susceptible to the very reporting bias the experimenter was trying to avoid.

In the first version of the experiment, the researcher decided which participant would play which role based on the flip of a coin, and then simply moved between the two adjoining laboratories to communicate with the individuals as required. Doing this in the second version of the study would alert the participants to one another's presence, and in doing so would make it more likely that the experimental aims may be predicted. Guessing patterns and strategies may then again become an artefact, and so changing this element of the experiment was trialled. Instead of deciding the role of the participants upon arrival, the researcher staggered the appointments to increase the chances that the participants would arrive separately. The first to reach the laboratory would be instantly taken into the room, and how and when they should provide surveillance would be explained to them. They were given chance to practice and to ask any questions they needed to in the hope that this should negate any need to revisit them during the experiment as this may draw attention to their presence.

The second person to arrive played the role of the observed participant, and their counterpart had been asked to stay as still and as quiet in the adjacent laboratory as possible whilst the protocol was explained to avoid conventional detection. Whilst the experiment was running, the participant who was to be observed was again asked to wear their headphones and so this risk was substantially diminished. Fortunately, as well as the internal door, both halves of the laboratory featured external access. This meant that if communication was unavoidable with the participant next door, the experimenter would always leave one laboratory via its exit to re-enter the adjacent room via the corridor so that it appeared as though they were leaving the participant entirely to give them privacy, rather than visiting another. Whilst more laborious for the experimenter, this level of attention to detail was worthwhile as subsequent interviews with the pilot study participants revealed that none were aware of their counterpart in the adjacent room, and none guessed that surveillance was taking place.

Another important change was that whilst the order of the tasks was randomised in Study One to avoid practice effects accounting for possible positive results, in Study Two the order was static. This was because for the reporting bias discovered in Study One to be eradicated, it was essential that the participants complete the passive tasks before undertaking the self-report experiment as this could highlight that surveillance was taking place.

3.2.7 Pilot Study Three (Running the Stroop Test with EDA).

The results of the Study One task performance caused the researcher to theorise that it mattered to the participants if they felt they were being tested, and so decided that for the second study, participants' EDA would be monitored whilst they were taking the Stroop test to investigate whether this 'pressure' would activate their EDA response to surveillance. Whilst the experimenter deemed this change to be a methodological improvement, its practicality was questioned as the Nexus EDA equipment relies on the relative stillness of the person it is monitoring, whilst the Stroop test requires the pressing of keys on a keyboard. For this reason, a pilot test was conducted to determine whether the small movements required might account for changes in the EDA readings. These tests showed that by applying the EDA monitor to the participant's non-dominant hand, and having them keep still except for their dominant hand, which was used respond to the Stroop test, the EDA readings were not affected.

It is also worth noting that even if the increased movement had heightened the EDA readings generally, the researcher was not looking for differences in EDA means between studies, but was specifically looking for differences between when the participants were (and were not) under surveillance. An overall consistent change would not therefore have presented a problem so long as differences between conditions would be valid. Tests showed that EDA readings were reliable in all participants who followed the suggested protocol, and the necessity to make clear the importance of stillness was passed onto all subsequent participants.

3.2.8 Procedure.

Having already conducted a focus group to consider the ethical considerations necessary for deceptive research prior to Study One, the experimenter briefed participants, invited and answered their questions, obtained their consent (see Appendix G), reassured them regarding their confidentially, informed them that that they can withdraw themselves and their data at any point, and debriefed them fully (see Appendix H) in accordance with the standard British Psychological Society (BPS) Code of Ethics and Conduct.

In the first study, participants could apply in pairs, or as individuals to be paired with another participant by the experimenter. This however proved problematic with consistent late arrivals and time lost through waiting for participants who never kept their appointment, and so this time participants were recruited singularly. Upon arrival, participants were offered a seat at a computer and were instructed to read and complete the consent form issued to them, and were offered the chance to ask any questions they may have. All who applied subsequently consented to take part. They were then invited to complete four questionnaires; the AEI, the O-LIFE, the CPES/TLS, and a short demographic questionnaire. All questionnaires were again completed online using the Qualtrics survey programme to allow the participant to enter their responses directly into the database to eradicate possible errors in transfer. The experimenter was present, but was never directly observing to avoid participants' performance anxiety, or demand characteristics. Once the observed participant had completed the questionnaires, the tasks they needed to complete were explained to them, but this time without detailing that surveillance would be taking place whilst they did so.

3.2.9 Study Two, Protocol One (Measured by EDA - With and Without Stroop)

The EDA protocol was divided into two distinct halves. The first half was an exact replication of EDA element of Study One (see section 2.2.11) to investigate its apparent lack of positive results - with the only difference being that the participant did not know that surveillance would take place this time. However, the second half was identical in every way to the first half except that the observed participant completed a pre-loaded computerised version of the Stroop test (Ridley, 1935) on the screen in front of them *whilst* their EDA was measured.

The basic principles of the Stroop test were explained to them, and it was made clear that they should press the letter on the keypad that corresponded with the colour that each word which appeared in front of them was written in, rather than what it actually said. A short practice session was held in the experimenter's presence to ensure the task was understood, and to ensure that the participant could undertake the task whilst remaining still enough to not adversely affect the EDA readings through movement alone. All participants were able to do so except one, and their experiment was abandoned at this stage.

It was explained to each participant that they will take the Stroop test four times, but they were not told that they would be a) under no surveillance at all, b) watched, b) listened to, c) watched and listened to until *after* their participation was complete. The participant was then instructed to wear the headphones, but it was not made clear to them that this was to ensure against them hearing movements in the adjacent room, so as to not alert them to the fact that observation may occur. The experimenter also left by the exit door rather than through the door which led to the adjacent laboratory to further enhance the idea that they would not be present during the experiment.

3.2.10 Study Two, Protocol Two (Measured by Task Performance)

This experiment was identical in every way to its equivalent in Study One (see section 2.2.12), except that this time it was conducted with the EDA equipment in place and attached to the participant to measure the electrical conductance of their skin to indicate physiological arousal whilst taking the test, and the participant was not aware of the randomised periods of surveillance.

3.2.11 Study Two, Protocol Three (Measured by Self-Reports)

The observed participant remained in the observation lab and sat at a computer with the PsycLab software pre-loaded. They were instructed to read the text that would appear on the monitor in front of them for 20-seconds until the programme asked via the screen whether they felt they had been watched, at which time they were to indicate their answer on the keypad (see section 2.2.10 for further details). Similarly, the programme then asked them to register whether they felt they had been listened to and the participant answered in the same way. This time though it was not highlighted to them that they would watched or listened to. It was however made clear that the task would last four consecutive minutes.

As in Study One, the observed participant was instructed to start reading the words aloud immediately upon the programme starting - but this was redesigned for Study Two with an increased delay at the beginning to ensure the experimenter had enough time to leave by the exit door, but enter the adjacent observation laboratory unbeknown to the participant so that they were not alerted to possible observation. The programme was then run twice to double the amount of trials each participant completed in Study One, with the intention of increasing statistical power.

Table 13

The Sequence of Events for Participant and Experimenter

		Measurement of surveillance detection (static order)		
Participant 1	Answers online questionnaires	EDA (5 mins)	EDA with task performance (4 x 1 mins)	Self-report (4 mins)
Participant 2	Preparing randomisation sequences	Provide surveillance	Provide surveillance	Provide surveillance

To check for ecological validity, participants were asked upon completion of the experiment, but before they received their debrief forms whether they suspected any elements of the study had not been disclosed to them. This was intended to protect the integrity of the findings. The debrief form was then issued, and participants were reminded of their freedom to withdraw their data, and that they should seek support if they found any elements of the study unsettling. Whilst the importance of not relaying details of the experiment to anyone else was specifically mentioned in the debrief form, this was also stressed verbally before the participants left the laboratory. This was the case in Study One due to the experiments' dependency on future participants' unfamiliarity with the experimental design, but it was even more important in Study Two due to the need for the information regarding surveillance taking place to be delayed until after the experiment had concluded.

Participants' psychophysiological responses, differences in their task performance, and their self-reported awareness of surveillance under the various conditions were subsequently compared against whether and when they actually *were* under surveillance to investigate possible relationships and patterns. Whether the psychosocial and neurological factors under investigation may be linked to such abilities were also analysed via one-way ANOVAs, one-sample t-tests, paired-samples t-tests, Pearson's correlation analysis, and multiple regressions.

3.3 STUDY TWO RESULTS

3.3.1 Treatment of the Imperfect EDA Data.

To determine whether changes in participants' EDA can be used to measure surveillance detection, participants' mean EDA values (in micro Siemens) under the four surveillance conditions (no surveillance, listened to, watched, watched and listened to) were calculated. The descriptive statistics (see Appendix AA) for the EDA measure can be seen in table 14 below. It should be noted that EDA mean scores of participant 33 in the 'listened to' condition (233.78) and the 'no surveillance' (44.78) condition were deemed to be outliers and so were removed.

Table 14

Mean EDA Descriptive Statistics

Mean EDA score	Mean micro Siemens of both 30 second epochs combined	Std. Deviation
No surveillance	5.44	2.81
Listened to	5.44	2.79
Watched	5.42	2.85
Watched and listened to	5.42	2.88

3.3.2 Analysing the EDA Scores Using a One-Way ANOVA.

A one-way repeated measures ANOVA (see Appendix BB) was conducted to determine whether there was a statistically significant difference in EDA means across conditions (none, watched, listened to and both). Mauchly's tests of sphericity indicated that the assumption of sphericity was violated, $X^2(5) = 56.72$, $p = .00023$. Epsilon was 0.76 as calculated according to Greenhouse and Geiser (1959), and was used to correct the repeated measures ANOVA. There was no significant difference in EDA means across conditions $F(2.27, 244.58) = .006$, $p = .997$, partial $w^2 = .000031$.

3.3.3 Transforming the Initial EDA Values Into EDA_d Values.

As in Study One, to take into account that the EDA means do not consider individual variances in EDA readings, new EDA figures (EDA_d) were calculated by deducting each participants' 'no surveillance' value from their surveillance (listened to, watched, watched and listened to) values. The mean and standard deviation for these recalculated EDA_d values can be seen in the below table (see Appendix CC).

Table 15

EDA_d Descriptive Statistics

Recalculated EDA values	Mean	Std. Deviation
Listened to	1.72	18.03
Watched	-0.37	4.02
Watched and listened to	-0.37	3.99

3.3.4 Correlations Analysis of Individual Differences (EDA_d Values).

Participants' EDA_d scores related to Study Two were analysed for relationships with each of the individual differences (see Appendix DD for descriptive statistics) under investigation to examine any meaningful patterns or trends. However, none of the other variables under investigation were found to be statistically significant using Pearson's correlation co-efficient analysis (see Appendix EE).

3.3.5 Regression Analysis of Individual Differences (EDA_d Values).

The association of individual differences to EDA ratio score were explored via three multiple regressions (see Appendix FF), with the observed variable conditions; (watched, listened to and both) as outcome variables and the nine individual differences measures (AEI experience [x2], AEI belief [x2], Unusual Experiences, Cognitive Disorganisation, Introvertive Anhedonia, Impulsive Nonconformity, and CPES) as predictor variables. All assumptions for multiple regression (linearity, non-multicollinearity, residual independence and normal distribution, homoscedasticity, absence of influential cases, and sufficient power) were met. Of the three multiple regressions, all returned non-significant models.

3.4 EDA With Stroop Test Measure

3.4.1 Analysing the Results of the EDA With Stroop Test Measure.

Following the results of Study One in which participants performing the Stroop test showed that their decision making seemed to be affected when they were being watched, and when they were being watched and listened to - the experimenter theorised that the stress or pressure of being tested created a 'need' to detect surveillance - and that this in turn may have enabled the ability to do so. The experimenter subsequently decided to run a version of the EDA element of the experiment *whilst* the participants were taking the Stroop task in an attempt to test this theory. To determine whether changes in participants' EDA can be used to measure surveillance detection, participants' mean EDA values (in micro Siemens) under the four surveillance conditions (no surveillance, listened to, watched, watched and listened to) were again measured. The descriptive statistics (see Appendix GG) for the EDA with Stroop measure can be seen in table 16 below.

Table 16

Mean EDA with Stroop Test Descriptive Statistics

Mean EDA score	Mean	Std. Deviation
No surveillance	3.66	1.75
Listened to	3.78	1.60
Watched	4.04	1.55
Watched and listened to	3.93	1.57

3.4.2 Analysing the EDA With Stroop Scores Using a One-Way ANOVA.

A one-way repeated measures ANOVA (see Appendix HH) was conducted to determine whether there was a statistically significant difference in EDA Stroop scores across conditions (none, watched, listened to and both). Mauchly's tests of sphericity indicated that the assumption of sphericity was violated, $X^2(5) = 57.08$, $p = 000227$. Epsilon was 0.74 as calculated according to Greenhouse and Geiser (1959), and was used to correct the repeated measures ANOVA. There was a statistically significant difference in EDA Stroop scores across conditions $F(2.23, 238.45) = 8.10$, $p = .00022$, partial $w^2 = .07$.

3.4.3 Analysing the EDA With Stroop Scores Using Paired-Samples T-Tests.

Paired-samples t-tests were conducted on to compare the mean EDA scores associated with the surveillance conditions (listened to, watched, watched and listened to) with the control condition (no surveillance) to investigate whether they were statistically different from one another. The results (see Appendix II) follow below.

PAIRING 1: No surveillance / listened to. A paired-samples t-test was conducted to compare participants' mean EDA when taking the Stroop Test under no surveillance, against their mean EDA when taking the Stroop Test whilst being listened to. There was no significant difference in their mean EDA scores between these two conditions; $t(107) = -1.564$, $p = 0.121$ (two-tailed), $d = .07$ (-0.19, 0.34).

PAIRING 2: No surveillance / watched. A paired-samples t-test was conducted to compare participants' mean EDA when taking the Stroop Test under no surveillance, against their mean EDA when taking the Stroop Test whilst being watched. Participants took significantly longer to complete the Stroop Test whilst being remotely watched, than when they were under no surveillance; $t(107) = -3.761$, $p = 0.001$ (two-tailed), $d = .23$ (-0.04, 0.50).

PAIRING 3: No surveillance / watched and listened to. A paired-samples t-test was conducted to compare participants' mean EDA when taking the Stroop Test under no surveillance, against their mean EDA when taking the Stroop Test whilst being watched and listened to. Participants took significantly longer to complete the Stroop Test whilst being remotely watched and listened to, than when they were under no surveillance; $t(107) = -2.668$, $p = 0.009$ (two-tailed), $d = .16$ (-0.10, 0.43).

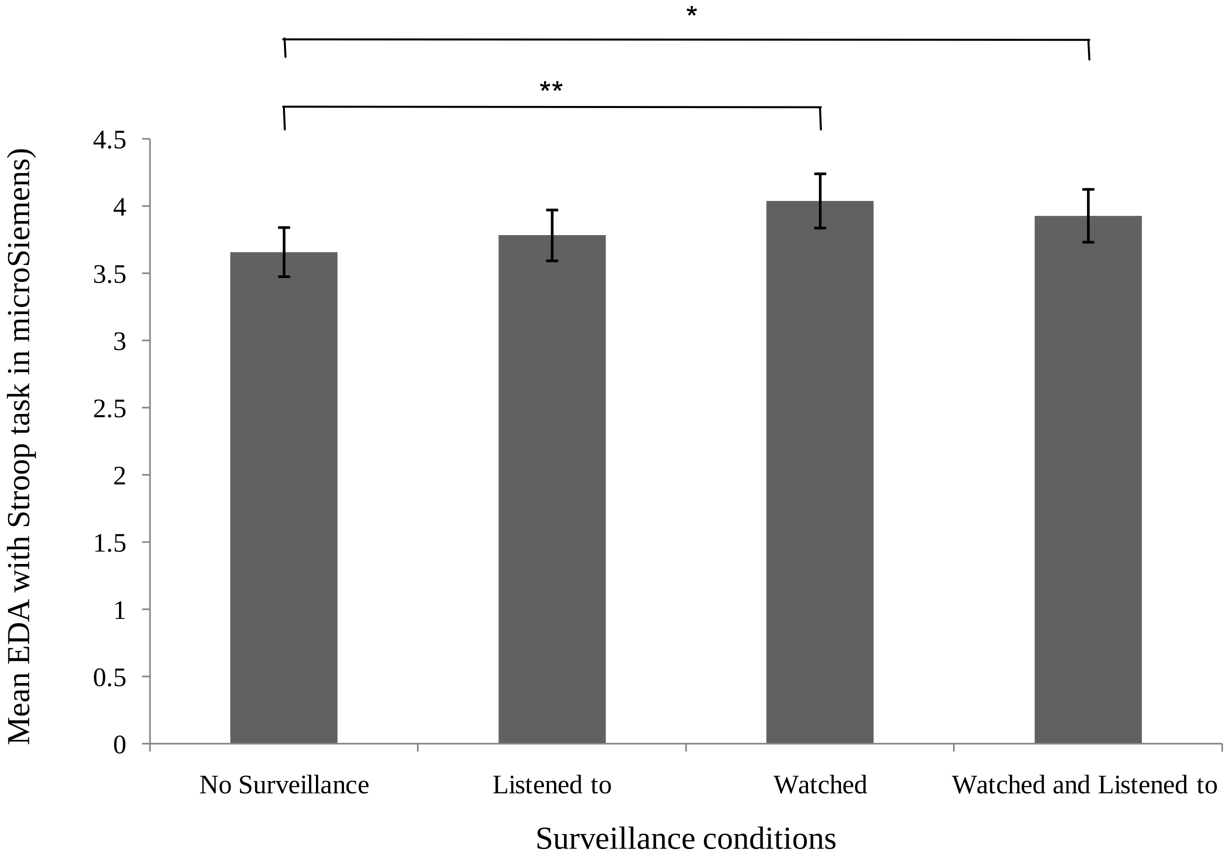


Figure 3 Mean EDA with Stroop task scores compared against 'no surveillance'.

Note. Paired-tests revealed that two of the EDA values were significantly different to the control group.

* One-sample t-test significant at the 0.05 level (two-tailed).

** One-sample t-test significant at the 0.01 level (two-tailed).

3.4.4 Transforming the Initial EDA Stroop Values Into EDA_d Stroop Values.

As in Study One, to take into account that the EDA means do not consider individual variances in EDA readings, new EDA subtracted differences Stroop figures (EDA_d Stroop) were calculated by deducting each participants' 'no surveillance' value from their surveillance (listened to, watched, watched and listened to) values. The mean and standard deviation for these recalculated EDA_d Stroop values can be seen in the below table (see Appendix JJ).

Table 17

EDA_d Stroop Descriptive Statistics

Recalculated EDA_d values	Mean	Std. Deviation
Listened to	.1242	.82531
Watched	.3808	1.05235
Watched and listened to	.2703	1.05270

3.4.5 Correlation Analysis of Individual Differences (EDA_d With Stroop).

Participants' EDA_d with Stroop scores related to Study Two were analysed for relationships with each of the individual differences (see Appendix DD for descriptive statistics) under investigation to examine any meaningful patterns or trends. However, none of the variables under investigation were found to be statistically significant using Pearson's correlation co-efficient analysis (see Appendix KK).

3.4.6 Regression Analysis of Individual Differences (EDA_d With Stroop).

The association of individual differences to EDA_d Stroop scores were explored via three multiple regressions (see Appendix LL), with the observed variable conditions; (watched, listened to and both) as outcome variables and the nine individual differences measures (AEI experience [x2], AEI belief [x2], Unusual Experiences, Cognitive Disorganisation, Introvertive Anhedonia, Impulsive Nonconformity, and CPES) as predictor variables. All assumptions for multiple regression (linearity, non-multicollinearity, residual independence and normal distribution, homoscedasticity, absence of influential cases, and sufficient power) were met. Of the three multiple regressions, all returned non-significant models.

3.5 Study Two, Protocol Two (Detection by Task Performance)

3.5.1 Analysing the Results of the Task Performance Measure.

In an attempt to replicate the significant findings related to Task Performance in Study One, whether participants' cognitive task performance would again be affected by covert surveillance was tested by them taking an online Stroop Test (Ridley, 1935) under the four surveillance conditions (no surveillance, listened to, watched, watched and listened to). The participants again showed a high level of attention, as none made any mistakes. Therefore, only the response time was used in the Stroop Test performance analysis. The descriptive statistics (see Appendix MM) of the time taken for participants to complete the Stroop Test can be seen in the table below.

Table 18

Task Performance Descriptive Statistics (Measured in Milliseconds)

Task performance conditions	Mean	Std. Deviation
No surveillance	14510.17	2284.34
Listened to	14548.04	2758.21
Watched	15020.78	3046.87
Watched and listened to	15180.56	3226.62

3.5.2 Analysing the Task Performance Scores Using a One-Way ANOVA.

A one-way repeated measures ANOVA (see Appendix NN) was conducted to determine whether there was a statistically significant difference in Stroop scores across conditions (none, watched, listened to and both). Mauchly's tests of sphericity indicated that the assumption of sphericity was violated, $X^2(5) = 15.25$, $p = .009$. Epsilon was 0.91 as calculated according to Greenhouse and Geiser (1959), and was used to correct the repeated measures ANOVA. There was a statistically significant difference in Stroop scores across conditions $F(2.74, 276.48) = 3.02$, $p = .035$, partial $w^2 = .029$.

3.5.3 Analysing the Task Performance Scores Using Paired-Samples T-Tests.

Paired-samples t-tests (see Appendix OO) were conducted to compare the surveillance conditions (listened to, watched, watched and listened to) with the control condition (no surveillance) to investigate whether there were significant differences in the time taken for participants to respond correctly when the word (red, green, blue etc.) appearing on their screen did not match the colour in which it was written.

The test data passed most of the assumptions required (see Appendix PP) for one-sample and paired samples t-tests, but the data was skewed. Due to the large sample size however, it was decided to still use t-tests rather than Wilcoxon rank sum analysis as with a sample of over 100 responses t-tests remain robust enough to be valid for any distribution (Lumley, Diehr, Emerson, & Chen, 2002).

PAIRING 1: No surveillance / listened to. A paired-samples t-test was conducted to compare how long participants took to complete the Stroop Test whilst they were under no surveillance, against how long it took them when they were being listened to. The difference in the time taken between these two conditions was not significant; $t(105) = -.049$, $p = 0.481$ (one-tailed), $d = .01$ (-0.25, 0.28).

PAIRING 2: No surveillance / watched. A paired-samples t-test was conducted to compare how long participants took to complete the Stroop Test whilst they were under no surveillance, against how long it took them when they were being watched. The difference in the time taken between these two conditions was not significant; $t(104) = -1.48$, $p = 0.072$ (one-tailed), $d = .19$ (-0.08, 0.45).

PAIRING 3: No surveillance / watched and listened to. A paired-samples t-test was conducted to compare how long participants took to complete the Stroop Test whilst they were under no surveillance, against how long it took them when they were being watched and listened to. The time taken for the participants to complete the task whilst being watched and listened to was significantly longer than when they were under no surveillance; $t(111) = -2.10$, $p = 0.019$ (one-tailed), $d = .24$ (-0.03, 0.51).

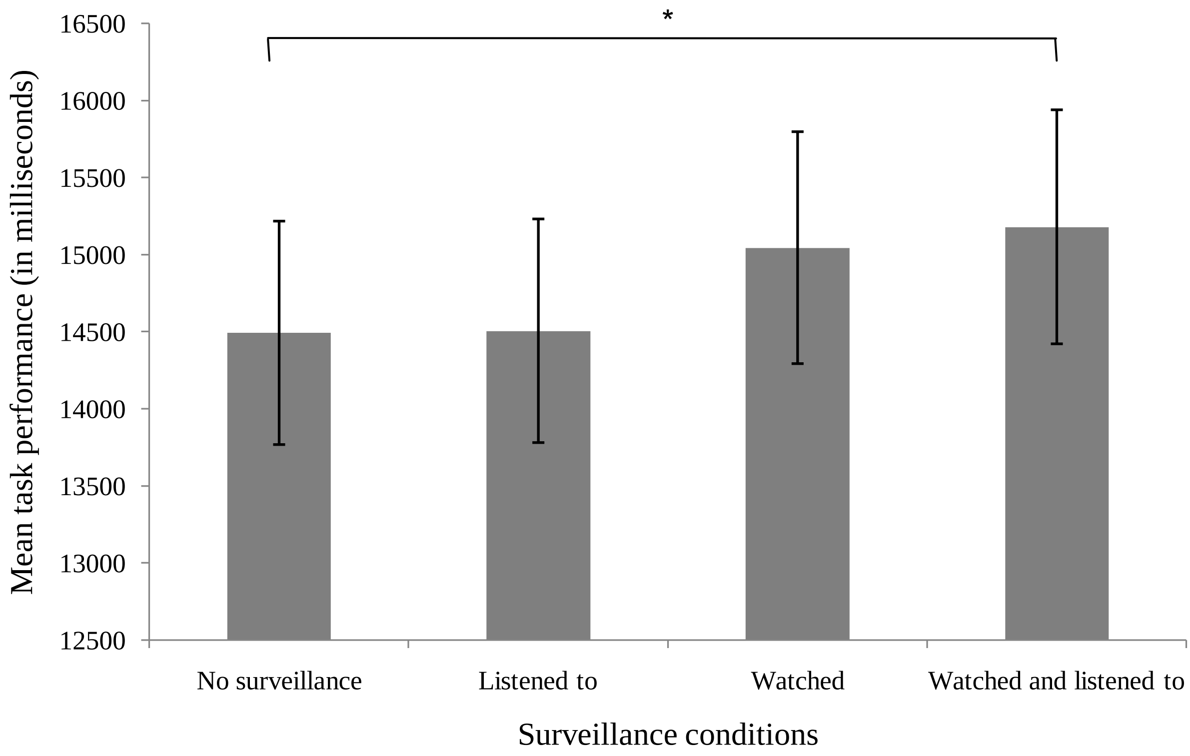


Figure 4 Significant task performance pairings.

Note. Only one of the combinations were significant, and this was at the 0.05 significance level.

* One-sample t-test significant at the 0.05 level (two-tailed).

** One-sample t-test significant at the 0.01 level (two-tailed).

3.5.4 Correlation Analysis of Individual Differences (Task Performance).

Participants' task performance scores related to Study Two were analysed for relationships with each of the individual differences (see Appendix DD for descriptive statistics) under investigation to examine any meaningful patterns or trends. None of the factors under investigation were found to be statistically significant using Pearson's correlation co-efficient analysis (see Appendix QQ).

3.5.5 Regression Analysis of Individual Differences (Task Performance).

The association of individual differences to Stroop test score were explored via four multiple regressions (see Appendix RR), with the observed variable conditions; (none, watched, listened to and both as outcome variables and the nine individual differences measures (AEI experience [x2], AEI belief [x2], Unusual Experiences, Cognitive Disorganisation, Introvertive Anhedonia, Impulsive Nonconformity, and CPES) as predictor variables. All assumptions for multiple regression (linearity, non-multicollinearity, residual independence and normal distribution, homoscedasticity, absence of influential cases, and sufficient power) were met. Of the four multiple regressions, all returned non-significant models.

3.6 Study Two, Protocol Three (Detection Measured by Self-Reports)

3.6.1 Treatment of the Imperfect Self-Report Data.

There were two missing data points within the self-report results which could be explained by the participants' failure to respond within the designated time (5 second) timeframe. These missing data points were replaced using the equivalent response for that participant, so if the participant had responded 'yes' when asked whether they felt they were being watched, but did not respond when asked the same question under the same condition later within the experiment, the researcher assumed the same response and entered it as such into the data set. Out of 3,520 possible self-report responses, only two were missed they did not occur within the same participants' data. Whilst the experimenter concedes that adding these data points created increased noise, it was decided that the analysis would become unnecessarily complicated if they were omitted. It was also felt that as the missing data points represented only 0.06% of the overall responses, completing the data would be unlikely to systematically affect the results, and that these participant's responses were too valuable to simply dismiss. Indeed, as before the replaced responses were divided equally between positive (YES) and negative (NO) - and so this method was the equivalent of the alternative option of randomly allocating the responses. In either event, the changes equate to zero whilst avoiding the need for recalculating all the probabilities.

3.6.2 Analysing the Self-Report Scores for Possible Bias.

The self-report element of Study One highlighted a general tendency for participants to over-report the feeling of being observed in trials, so the results from Experiment Two were examined for the same potential issue. Participants made 1,770 positive observation responses overall out of a possible 3,520, equivalent to a 50.29% response rate, which was not significantly higher than mean chance expectation (50%) where $t(109) = 1.67$, $p = 0.424$ (one-tailed). This indicates that the experiment redesign had its intended effect, and by not telling the participants that they would be randomly watched and listened to, the tendency over report surveillance was eradicated (see Appendix SS). There was however a significantly greater tendency to report being watched than being listened to, $t(109) = 2.041$, $p = 0.04$ (two-tailed), and the reported observations in each mode exhibited a significant positive correlation $r(109) = .465$ $p = .001$ indicating very little difference within-subjects to report one mode of observation over the other.

3.6.3 Analysing the Self-Report Scores Using Paired Samples T-Tests

The test data passed most of the assumptions required (see Appendix TT) for one-sample and paired samples t-tests, but the data was skewed. For ease of comparison with Experiment One's t-test results, using the same method of analysis was preferable. Despite the distribution of the data, this was valid due to the large sample size and the robustness of the t-test (Lumley et al., 2002).

Table 19

*Signal Detection Rates for the 'Listened to' Mode of Response for all Trials (Nt = 4) for all Participants (N = 110) for all of the Four Conditions **

	True	True %	False	False %
Positive	467 (hit)	53.07%	413 (false alarm)	46.94%
Negative	441 (correct rejection)	50.11%	439 (miss)	49.89%
Total		51.59%		48.42%

*[total number of trials across all conditions and all participants = 1792]

Whilst in the predicted direction, a paired t-test comparing correct Listened to self-report scores (Listen & Both conditions), $M = 4.25$, $SD = 2.18$, with incorrect Listened to self-report scores (Watched & No surveillance conditions), $M = 4.01$, $SD = 2.23$, was not significant $t(109) = 1.30$, $p = 0.10$ (one-tailed), $d = .11$ (-0.16, 0.37).

Table 20

Signal Detection Rates for the 'Watched' Mode of Response for all Trials (Nt = 4) for all Participants (N = 110) for all of the Conditions *

Watched	True	True %	False	False %
Positive	498 (hit)	56.6%	382 (false alarm)	43.4%
Negative	451 (correct rejection)	51.3%	429 (miss)	48.7%
Total		53.95%		46.05%

*[total number of trials across all conditions and all participants = 1792]

A paired t-test comparing correct Watched self-report scores (Watch & Both conditions), M = 4.53, SD = 1.97, with incorrect Watched self-report scores (Listen & No surveillance conditions), M = 3.90, SD = 1.97, showed a significant increase $t(109) = 2.87$, $p = 0.025$ (one-tailed), d = .32 (0.05, 0.59).

Table 21

Signal Detection Rates for the 'Watched and Listened to (Both and None)' Mode of Response for All Trials (Nt = 4) for all Participants (N = 110) for all of the Four Conditions *

	True	True %	False	False %
Positive** (MCE = 25%)	160 (hit)	36.36%	342 (false alarm)	25.91%
Negative*** (MCE = 25%)	160 (correct rejection)	36.36%	332 (miss)	25.15%
Mean		36.36%		25.53%

* [total number of trials across all conditions and all participants = 1760 (440 for true, 1320 for false]

** True/False Positives are those trials where participants responded that they had been both Watched and Listened to

*** True/False Negatives are those trials where participants responded that they had not been either Watched or Listened to

A paired t-test comparing correct combined both Watched and Listened To self-report scores ('Both' condition), *M* = 1.45, *SD* = 1.33 with incorrect combined both Watched and Listened To self-report scores (None condition), *M* = 1.03, *SD* = 1.23, showed a significant increase, *t*(109) = 2.82, *p* = .003 (one-tailed), *d* = .33 (0.06, 0.59).

3.6.4 Comparing Self-Report Scores With Chance Expectation.

To explore participants' tendency to either accurately or falsely detect when they were under surveillance, and when they were not under surveillance; two exploratory one-sample t-tests were performed to enable a comparison with the literature which tends to compare correct self-reports with chance (50%), and how participants' tendency to be correct is often greater than chance, whilst their tendency to be wrong is often lower than chance (Sheldrake, 2003a).

Table 22

*Signal Detection Rates (Red = Above Chance, Blue = Below Chance) for all Responses Combined, for all Trials (Nt = 4) for all Participants (N = 110) for all of the Four Conditions **

	True	True %	False	False %
Positive** N = 2,640	1285 (hit)	48.67%	1484 (false alarm)	56.21%
MCE 50%		41.67%		58.33%
Negative*** N = 2,640	1270 (correct rejection)	48.11%	1469 (miss)	55.64%
MCE 50%		41.67%		58.33%
Total		48.39%		55.93%
% above MCE		6.72%		-2.40%

* [total number of trials across all conditions and all participants = 1760 (440 for true, 1320 for false]

** True/False Positives are those trials where participants responded that they had been both Watched and Listened to

*** True/False Negatives are those trials where participants responded that they had not been either Watched or Listened to

Below are the one-sample t-tests (see appendix UU) that support the literature.

A one-sample t-test found that the total true response rate was significantly greater than mean chance expectation, $t(5279) = 9.77$, $p = 0.2 \times 10^{-21}$, $d = .13$ (.11, .16).

A one-sample t-test found that the total false response rate was significantly lower than mean chance expectation, $t(5279) = 3.52$, $p = 0.0004$, $d = -.048$ (-.075, -.021).

The above analyses would appear to support the tens of thousands of trials (Sheldrake, 2003) which produced statistically significant results showing that participants score above chance, thereby supporting the remote staring detection interpretation, and refuting suggestions of responses bias affecting positive outcomes.

3.6.5 Transforming the Self-Report Scores Into Ratio Scores.

Self-report ratio scores were again calculated by taking each of the self-report scores (no surveillance, listened to, watched, watched and listened to) and dividing each of these totals by the self-report total so that the ratio score is relative to the participants' overall reporting of surveillance detection to correct for the general tendency to over-report surveillance detection. Whilst this was less necessary for Study Two than it was for Study One due to the reporting bias issue being largely dealt with via the methodological change, these figures were used to ensure compatibility between studies so that whether the reporting bias issue had a tangible effect could be investigated. The descriptive statistics can be seen in the below table (see appendix VV).

Table 23

Recalculated Self-Report Ratio Descriptive Statistics

Self-report ratio conditions	Mean	Std. Deviation
No surveillance (N=2)	0.15	0.09
Listened to (N=4)	0.34	0.08
Watched (N=4)	0.34	0.10
Watched and listened to (N=2)	0.17	0.05

3.6.6 Correlation Analysis of Individual Differences (Self-Report Ratio Scores).

Participants' self-report ratio scores related to Study Two were analysed for relationships with each of the individual differences (see Appendix DD for descriptive statistics) under investigation to examine any meaningful patterns or trends. The below are the individual differences variables found to be statistically significant using Pearson's correlation co-efficient analysis (see Appendix WW). It should be noted though, that none of these remained significant once a Bonferroni correction was applied.

Table 24

Correlations Between Participants' Individual Differences Variables and Self-Report Ratios

N = 110	No surveillance (*M* = .1455, *SD* = .08744	Listened to (*M* = .3374, *SD* = .07602)	Watched (*M* = .3437 *SD* = .09515)	Watched and listened to (*M* = .1731, *SD* = .05445)
AEI Experience (*M* = .3320, *SD* = .30361)	-.236	-.232	.410*	-.315
AEI Special Experience (*M* = .3522, *SD* = .29736)	-.215	-.234	.406*	-.326
AEI Belief (*M* = .4721, *SD* = .32760)	-.051	-.146	.263	-.308
AEI Special Belief (*M* = .5185, *SD* = .30762)	-.032	-.159	.261	-.308

Unusual Experiences (M = 4.83, SD = 2.851)	.105	-.162	.063	.015
Cognitive Disorganisation (M = 4.39, SD = 2.603)	-.010	-.174	.145	-.068
Introvertive Anhedonia (M = 4.94, SD = 1.473)	.128	.175	-.242	.138
Impulsive Non-Conformity (M = 4.66, SD = 2.038)	.330	-.008	-.297	.357
CPES (M = 5.02, SD = 3.543)	.352	.003	-.165	-004

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

NOTE: None of the significant correlations remained so after the Bonferroni correction (alpha = .0055).

3.6.7 Regression Analysis of Individual Differences (Self-Report Ratio Scores).

The association of individual differences to self-report ratio scores were explored via four multiple regressions, with the observed variable conditions; (none, listened, watched, both) as outcome variables and the nine individual differences measures (AEI experience [x2], AEI belief [x2], Unusual Experiences, Cognitive Disorganisation, Introvertive Anhedonia,Impulsive Nonconformity, and CPES) as predictor variables. All assumptions for multiple regression (linearity, non-multicolinearity, residual independence and normal distribution, homoscedasticity, absence of influential cases, and sufficient power) were met. Of the four multiple regressions none returned significant models (Appendix XX).

3.7 DISCUSSION OF STUDY TWO

Amended hypotheses were developed for Study Two to build upon what was discovered following analysis and consideration of Study One's results, and the methodological changes this inspired. Hypotheses regarding covert surveillance detection accuracy were as follows; EDA arousal levels would be significantly different when participants were under covert surveillance compared to when they were not if they were taking the Stroop test, their Stroop Test completion times would be significantly different under surveillance conditions compared to the control condition, and accurate self-reporting of covert surveillance would be significantly greater than both chance expectation and inaccurate self-reports. In relation to the individual differences, it was hypothesised that participants' ability to detect covert surveillance will be correlated with the individual differences variables. Analysis of the data resulting from Study Two demonstrated that the majority of these hypotheses were supported and there were more significant findings than in Study One, suggesting that the methodological changes influenced the outcomes as predicted to some extent.

Whilst the results of the EDA protocol in Study One suggested that this method of measuring remote surveillance detection is not an effective means of doing so, it was repeated in Study Two to investigate whether this was due to the participants' expectation of surveillance. The only difference between the EDA protocol in the first and second study was that participants were not primed to expect that they would be watched and listened to - however, this methodological change had little effect on the results. A one-way repeated measures ANOVA showed that there were again no significant differences in participants EDA readings regardless of the surveillance conditions. These results appear to confirm that using participants' EDA alone is not a reliable or accurate measurement of surveillance detection - at least with the equipment and experimental methodology utilised in the current experiments.

Despite the lack of significant results obtained from Study One in relation to the EDA protocol, it remained somewhat surprising that the same was found following the partial replication in Study Two due to the strength of the evidence regarding this measure's effectiveness found in the literature. For example, Lobach and Bierman's (2004) findings suggested that it would be more likely that positive results would be obtained via using EDA to capture surveillance detection than via self-reports. Schmidt and Walach (2000) have warned however that this method is prone to artefacts, and this was demonstrated in Study One's EDA experiment when a standard deviation of 13 - 14% was found. This potential obstacle to detection was of concern as Lobach and Bierman, (2004) reported a seemingly genuine signal difference of only 1%. For this reason, even more care was taken in Study Two to emphasise the importance of staying still to all participants when the EDA protocol was replicated, and considerable effort was made to avoid any possible disturbances or distractions, however this made little difference to the results.

As with Study One, to take individual variances into account, EDA_d values were created to measure change within the participants EDA which was relative to their 'no surveillance' EDA. None of the variables under investigation were found to be statistically significant using Pearson's correlation co-efficient analysis, and none of the multiple regressions conducted returned non-significant models. Both analyses therefore indicate that the factors under investigation are not related to participants' remote observation abilities when measured by EDA alone.

Whilst the researcher had to consider that the results of the second version of the EDA protocol being similar to the first could mean that there is genuinely no remote surveillance effect to detect, analysis of the task performance protocol in Study One suggested that a feasible explanation may be found in participants 'need' to detect the surveillance, especially as EDA is a direct response to stress (Silverthorn et al., 2009). It was surmised that perhaps the lack of 'threat' or any 'necessity' to detect surveillance relating to the EDA element of the current experiment up until that point may have rendered this measurement redundant. Put another way, the absence of actual danger, or a reasonable and realistic substitute may be the reason participants' physiology was not responding to them being watched or listened to. This theory is supported by Sheldrake's (2003a) anecdotal evidence for scopaesthesia offered by Special Forces operatives, snipers, security guards and police officers who claim the ability to know when they are being watched is both real, and essential to their work. The common theme between these situations is that survival may depend on the awareness of being the focus of another's' attention.

The experimenter had previously considered that the safety of a laboratory based experiment in which considerable efforts had been made to make the participants feel relaxed may have actually been an obstacle to using EDA as a measure of detection, and could have effectively 'switched off' any extrasensory abilities the participants may have had due to the safety of their situation. Whilst the researcher accepts the impossibility of creating life-threatening and dangerous situations in which their experiments may be more likely to produce significant results should extrasensory ability exist, the possibility that the reason the task performance protocol in Study One produced significant results may be due to the stress involved in being tested was considered. To investigate this, the EDA protocol was again conducted - but this time the participants EDA was recorded whilst undertaking the same Stroop Test that was used in the Task Performance protocol previously.

A one-way repeated measures ANOVA was conducted to determine whether there was a statistically significant difference in participants' EDA during the Stroop task, and this was shown to be the case. When subsequent paired-samples t-tests were conducted to compare the mean EDA scores associated with the surveillance conditions against the control condition, two of the three combinations showed significant differences. It was revealed that participants' mean EDA scores whilst watched were significantly higher than when they were not under any surveillance, providing support for scopaesthesia. Additionally, participants' mean EDA scores whilst watched and listened to were significantly increased when compared to when neither was occurring - this not only lends weight to the proposed existence of the sense of being watched, but also to the experimenter's proposal that the sense of being listened to may enhance individuals surveillance detection ability. This finding would be noteworthy in and of itself, but when combined with the Task Performance findings of Study One in which the doubling-up of being watching and listened to provided the study's most significant result, it is even more interesting.

Comparing the lack of significant outcomes when examining the EDA scores as an independent measure (see section 3.3.2) with the significant results found when the Stroop Test was included (see section 3.4.3) could be considered compelling evidence that stress and/or threat is an important element of individuals' ability to remotely detect surveillance. The introduction of the Stroop Test to an otherwise identical protocol resulted in two of the three t-tests conducted being statistically significant, compared with when none of them were significant without the inclusion of the test. Whilst this methodological change was based on a logical leap, albeit inspired by the evidence extracted from the data analysed following Study One - the researcher cautiously suggests that this supports the idea that surveillance detection is to some extent determined by the *necessity* to detect. In other words, by simply making it 'matter' whether the participants were being watched through the introduction of a test element (and therefore adding stress by association), the change appears to have created a statistically significant difference to the participants' EDA scores, thus making this a seemingly accurate method of surveillance detection under the correct circumstances.

Following such encouraging results when the Stroop Test was included as part of the EDA protocol, the experimenter was keen to see if the individual differences under investigation would be shown to have significant relationships with differences in participants' EDA scores. To ensure fair comparison, and to again take into account that the EDA means do not consider individual variances - EDA_d figures were calculated by deducting each participants' 'no surveillance' value from their surveillance (listened to, watched, watched and listened to) EDA readings. This time however, the calculation was performed on the data from when the participants were having their EDA measured whilst taking the Stroop Test. Participants' EDA_d with Stroop scores related to Study Two were subsequently analysed for relationships with each of the individual differences under investigation to examine any meaningful patterns or trends. However, none of the variables under investigation were found to be statistically significant using Pearson's correlation co-efficient analysis despite the methodological change. Similarly, the association of individual differences to EDA_d with Stroop scores were also explored via multiple regressions, however all three returned non-significant models.

With the introduction of the Stroop Test having been shown to make a significant difference to participants' EDA scores, the experimenter next looked at whether the time participants took to complete this test was influenced by the surveillance conditions. Despite the comparative success of this element in Study One when the participants' decision making appeared to be significantly delayed by the 'watched' and the 'watched and listened to' conditions when compared to when they were under no surveillance - it was deemed that supporting evidence was needed and further research regarding task performance as a detection method was required.

Had the protocol in Study Two been an exact replication, the hypothesis would have been one-tailed due to the previous positive task performance results - however as the participants' knowledge that surveillance would take place was removed, this single, yet potentially important methodological change meant that the hypothesis remained two-tailed. This decision was taken after much consideration as the original hypothesis related to Task Performance was to some extent exploratory due to the comparative lack of previous research to support it, and the experimenter concluded that it was possible that the participants' knowledge that they would be watched and listened to could have caused the stress that may have led to the positive results. Indeed this idea seems to have been supported by the results, which were partially repeated in Study Two. A one-way repeated measures ANOVA showed there were statistically significant differences in the time participants took to complete the Stroop Test depending on which surveillance condition was taking place, and follow-up paired-samples t-tests revealed that whilst there was no significance difference in the time participants' took to complete the Stroop Test whilst being watched this time; they did take longer to do so when being watched and listened to. These findings could be argued not only to support scopaesthesia, and that the sense of being stared at is enhanced when coupled with the sense of being listened to; but also the experimenter's prediction that the effect would be weakened once participants did not know that surveillance would be taking place, and thus stress was reduced - however this is somewhat speculative without further research.

The literature supports the idea that being tested would increase stress, as when forced into completing non-voluntary exercises such as examinations or tests, neural stress-induced responses to pressure such as an upregulation of dynorphin/kappa opioid receptor is thought to underlie the negative affect associated with increased stress (Varlinskaya et al., 2020). Other researchers argue that the hyperexcitation of the amygdala (due to chronic stress) leads to stress-related responses such as a chronic itch (Pavlenko & Akiyama, 2019), which indicates that high-stakes or high pressure situations such as testing incites neural responses that may lead to stress-induced behaviour or negative emotions associated with stress.

Jamieson et al., (2020) argue that high-pressure testing environments influence changes in neuroendocrine levels (both cortisol and testosterone) and performance outcomes due to the mediating power of stress appraisals, which further supports the role of neural and neuro-chemical factors as mediators of affective stress and stress-induced behavioural responses. Real-world applications such as cognitive reappraisal highlights the role biological and neurological factors play in inducing stress responses, whereby cognitive reappraisal reduces cortisol reactivity to stress and increases cardiac flexibility (Jentsch & Wolf, 2020). Therefore, reducing physiological and neurological responses to stress appraisals also helps to reduce stress and stress-related responses when engaging in high pressure situations such as testing (Gnam et al., 2019) which highlights the importance of physiological and neurological factors as potential mediators of increased stress.

Stress may have been exaggerated in Study One due to the participants' knowing that they would be watched and listened to. The discovery that covert surveillance remained detectable even without this knowledge, albeit to a lesser extent in Study Two could be considered compelling evidence of the existence of scopaesthesia, and may even support acoustathesia - although it appears the latter may be undetectable unless coupled with the participant being watched.

The experimenters' decision to include a behavioural measure such as task performance was related to peoples' tendency to behave and perform differently when alone, to when they are in the presence of others (Griffin & Kent, 1998). The subsequent laboratory experiments tested the logic that if people perform tasks differently when under conventional observation, then they may do so when the observation is covert if they are detecting its presence via unconventional means. To some extent this prediction has been supported - but with the caveat that a stressful situation appears to be required for remote detection. Whilst Sheldrake's (2003) work suggested an effect was likely, his research was under very different circumstances. Indeed, the significant effects of Study One's results related to task performance was a surprise to the current experimenter, as at the time it had been found that self-reports and EDA measurements were not statistically significant means of detection, despite them being thought to be better signals of scopaesthesia. This made it theoretically unlikely that task performance could be a means of detection, as social facilitation dictates that individuals respond to the knowledge they are being observed by changing their behaviour. It logically follows that if the self-report measure demonstrated that the participants were not consciously aware of surveillance, and the EDA measure suggested they are not unconsciously aware - then task performance should also not be affected by covert surveillance.

If the experimenter is correct in deducing that it was indeed the introduction of stress which had such an effect, and that by making it of concern to the participants whether or not they were under surveillance resulted in them 'engaging' their extrasensory ability - then by not telling them that surveillance would occur in the second version of the task performance experiment, the researcher may have inadvertently 'switched off' the participants need to detect surveillance. This would explain why all but the theoretically most opposite surveillance conditions combinations of 'no surveillance' (when no attention was directed at participants) and 'watched and listened to' (when the most attention was given to the participants) was the only comparison in which the two means were consistently significantly different to each other. The decision to withhold the information that participants would be under surveillance was primarily aimed at resolving the reporting bias that was an issue in the self-report experiments, and so its effect on task performance was to some extent accidental. Despite this, the discovery is potentially important not only to research on this topic, but indeed to all research as it offers a valuable contribution to the subject of priming and expectation.

Having found the measure of task performance to be an accurate indicator of detection in both studies, the experimenter was keen to see if these findings shared a relationship with the individual differences variables. To take into account that task performance values do not consider individual variances, the data was again transformed to allow for this. The resulting EDA_d scores were analysed for relationships with each of the individual differences under investigation to examine any meaningful patterns or trends, however, there were no significant models found when the task performance completion times were explored via multiple regressions for Study Two. This was despite Introvertive Anhedonia and CPES being significant predictors of task performance while watched and listened to, and Impulsive Nonconformity being a significant predictor of task performance while under no surveillance in Study One.

The researcher considered whether the absence of any significant relationship with the schizotypy subscale of Impulsive Nonconformity could be linked to the experimenter not priming the participants to expect surveillance. The author's earlier work (Friday & Luke, 2014) demonstrated that Impulsive Nonconformity appeared to lead participants who scored highly on this negative trait to report *against* expectations. It was therefore tentively suggested that participants scoring highly on the Impulsive Nonconformity scale may have deliberately reacted in the opposite direction to which they were expected to - however in this second version of the experiment the researcher's expectation was not clear, and therefore high Impulsive Non-conformity scorers were not able to behave in this way. Harder to explain however is the lack of relationship between temporal lobe lability (CPES) and task performance as this was shown to be strongly and positively correlated in the first task performance experiment in the 'watched and listened to' category. This was predicted due to the research of Horowitz and Adams (1970), Luke et al. (2013) and Persinger et al. (2000); however, this relationship was apparently absent in Study Two and so appears not to be a robust correlate or predictor of remote observation detection.

With the EDA and Task Performance protocols analysed, the experimenter's attention turned to the self-report measure that is well supported by the literature (Colwell et al,. 2000; Radin, 2004;

Sheldrake, 1998, 1999, 2000, 2001, 2003, 2005b), and therefore was predicted to produce positive results in Study One. Whilst this method of covert surveillance detection failed to produce significant differences between surveillance conditions in the first experiment, and none of the individual differences variables remained significantly correlated after Bonferroni corrections were applied - it was considered that the methodological changes applied to the second experiment may result in different outcomes. By doubling the number of trials to increase statistical power, and by not announcing to the participants that surveillance would be taking place - it was hypothesised that the former would mean the results would become significant rather than just showing trends in the predicted direction, and that the latter would eradicate the reporting bias that was evident in the first experiment.

Analysis of the self-report scores in Study One highlighted a general tendency for participants to report the feeling of being observed during trials, so the results from the comparable experiment in Study Two were examined for the same potential issue. Participants made 1,770 positive observation responses overall out of a possible 3,520, equivalent to a 50.29% response rate. This was not significantly higher than mean chance expectation (50%) which indicates that the changes to the methodology had the intended effect - by not drawing the participants' attention to the fact that they would be periodically observed and/or listened to, their tendency to over report this due to expectation was dealt with overall.

Self-report ratio scores were again calculated by taking each of the self-report scores (no surveillance, listened to, watched, watched and listened to) and dividing these by the self-report total so that the ratio score is relative to the participants' overall reporting of surveillance detection. This was done for Study One to correct for the general tendency to over-report surveillance detection, and so was less necessary for Study Two as the reporting bias was dealt with via the methodological change - however these figures were used to ensure compatibility between studies.

Paired-samples t-tests were conducted to compare participants' correct self-reports with their incorrect self-reports for significant differences in scores for each of the surveillance conditions. When individuals' correct self-reports of being listened to were compared with their incorrect self-reports of being listened to - whilst in the predicted direction, the difference was not significant indicating that that any signal was either too weak to detect, or did not exist at all. This result would be expected to some extent as the literature on scopaesthesia suggests that whilst being watched is detectable, there is no research to suggest that the same is true of being listened to. However, this does mean that theoretically the participants' being watched should have produced independently significant results, and this was not the case in the current study. It is noteworthy though that research which employed a two-tailed hypothesis would have found this difference to be significant. This could be argued to indicate that a weak signal may be present, and to explain why previous studies have found significant results for this surveillance condition. Indeed, the figures within the watched condition were promising as the scopaesthesia related work of Colwell et al. (2000), Radin (2004) and Sheldrake (1998, 1999, 2000, 2001, 2003, 2005b) suggested a clear tendency for their participants to know when they were being watched, and in the current experiment, participants were correct 56.59% of the time in both the 'watched' and the 'watched and listened to' conditions - a finding in line with Sheldrake's (2003) repeatedly consistent correct response rate of around 55%.

A significant difference was found in the current study however when comparing correct combined 'watched and listened to' self-report scores with incorrect combined 'watched and listened to' self-report scores, providing further evidence that participants' remote surveillance detection ability is significantly enhanced if they are simultaneously watched and listened to. This would appear to make logical sense, as this is the most extreme of the surveillance conditions where the participants were receiving the maximum attention possible as they were being watched *and* listened to at the same time. Put another way, this evidence suggests that participants were able to detect being the focus of someone's complete attention, and could accurately and consciously discriminate between this and a complete lack of attention. The experimenter concedes though that this conclusion relies on 'complete attention' being defined as the observer's ability to see and hear their counterpart, and that there is no way of knowing the observers' intention or degree of focus. Subsequent one-sample t-tests comparing the total true response rate to mean chance expectation showed a significant increase as predicted, and the same was true when comparing the total false response rate to mean chance expectation.

It would therefore appear that the addition of more trials per participant to increase statistical power had a tangible effect, as did omitting to highlight that participants would be under surveillance until after the experiment had ended. These methodological amendments seem to have produced a significant result in relation to the 'watched and listened to' condition, with all other conditions showing results in the predicted direction. Indeed, when all correct responses were looked at in conjunction, the data showed that participants were correct 6.72% more often than would be expected by chance alone, which was highly significant and very much in line with the work of Sheldrake (1998, 1999, 2000, 2001, 2003, 2005b). This outcome should perhaps not be surprising as of the three detection measures used in the current study, there is more literature on self-reports than either psychophysiological or task performance. Indeed, analysis of tens of thousands of trials demonstrate an effect of around 5% above chance (Sheldrake, 2003) - similar to that found in the current study when looking at correct guesses overall. It is important to note that this effect appears to remain even after the important issues of sensory leakage and proper randomisation highlighted by researchers such as Marks and Colwell (2000), and the problem of reporting bias identified by Sheldrake (1998, 1999, 2000, 2001, 2005b) are addressed. Indeed the current study appears to demonstrate that whilst the practice and feedback regarding participants' accuracy may have a dramatic effect on participants' successful self-reporting as claimed by Colwell et al. (2000) it is not necessary for positive and convincing results of extrasensory surveillance detection.

Whilst the individual differences variables in Study One appeared to have no relationship with self-reports after the Bonferroni corrections were applied to take multiple analyses into account, and none of the multiple regressions returned significant models - this was perhaps not surprising when participants were shown to be unable to detect being watched or listened to, regardless of the surveillance conditions. Following the significant self-report results found in Study Two however, the experimenter thought it more likely that the individual differences under investigation may demonstrate trends or patterns in relation to self-reporting accuracy. Despite this, whilst both versions of the AEI Experience variable were found to be correlated with the 'watched' condition, this relationship did not survive Bonferroni correction for multiple analyses, and when the association of individual differences to self-report ratio accuracy were explored via multiple regressions, none returned significant models. These findings mean that whilst limited support was found for the hypothesis that correct self-reports will be significantly greater than incorrect self-reports when the number of trials per participant is doubled, and priming is avoided - this was not the case regarding the hypothesis that the individual differences variables will be correlated with remote surveillance detection ability, as this hypothesis was not met.

When the results of Study Two are looked at in combination, they seem to suggest that surveillance detection via extrasensory means may exist - but is elusive and difficult to capture, with partially replicated experiments producing very different results to those which did not include these seemingly small methodological amendments. The changes to the experiments from Study One to Study Two resulted in more significant outcomes related to self-reports and EDA response, however it resulted in less significant results related task performance. Despite this, there seems to be a consistency emerging from the results overall that suggests if the participants may be experiencing stress or anxiety, this improves their extrasensory surveillance detection accuracy.

With two lengthy and methodologically thorough studies demonstrating the possible existence of scopaesthesia even under well-controlled laboratory conditions, an obvious criticism of the research is its possible lack of ecological validity. The very environment which allowed the experimenter to eradicate the sensory leakage and the artefacts which have plagued many of the experiments detailed in the literature is the very reason the current studies have a low level of realism that can be applied to the real world.

The lack of any similarity between a laboratory environment, and events which usually take place in a real-world setting were always of concern to the researcher - especially as the findings have the potential to contribute to combating increasingly prevalent security issues being experienced internationally. As such, research which cannot be applied to improving the situation for individuals in actual scenarios which they are likely to experience has limited value, and so the researcher decided to test their emerging theory; namely that stress is important to the accuracy of covert surveillance detection, and that being listened to and watched in combination is the most detectable form. To address their concerns regarding the lack of ecological validity associated with non-laboratory settings, the researcher decided that the third study should consist of a field experiment that would further test their findings and the factors already under investigation.

Whilst field experiments promote high ecological validity and there is a diminished risk of demand characteristics if participants are not aware they are being studied, this research must be designed, conducted and analysed with its limitations in mind as this method of investigation is not without important limitations. Such experiments make it difficult to control for extraneous variables that could influence the results, create complications with obtaining fully informed consent from participants, and are less replicable and comparable than laboratory based experiments.

4 STUDY THREE.

4.1 Study Three Introduction

Having conducted Studies One and Two, the current research appears to support the 70% - 90% of the general public who believe it is possible to detect observation by extrasensory means (Braud et al., 1993a: Coover, 1913; Sheldrake, 2003; Thalbourne & Evans, 1992), and suggests that this belief could possibly be based on their genuine ability to do so. It also offers laboratory based evidence to the wide and varied body of research which led to the claim that observation detection accuracy levels would be approximately five percent higher than chance expectation (Sheldrake, 2003) which seemingly demonstrates that the phenomena of scopaesthesia may be both real, and measurable. However, the signal appears to be weak and difficult to detect.

The problems that experimenters have encountered with the detection of scopaesthesia, and indeed the laboratory based studies which failed to find any evidence of this ability may of course be due to its non-existence. However, it should also be considered that the mixed results, and the apparent weakness of signal may be due to other issues. Possibly the most widely accepted theoretical explanation for scopaesthesia is that such an ability would be evolutionarily advantageous (Sheldrake, 2003), and could have developed when humans were both hunters and the hunted, with this ever-present present danger leading to the development of the extrasensory ability to detect being watched.

If evolution were to offer an explanation for scopaesthesia either in full, or in part - then it may be naive to expect a sense which developed in response to threat and survival to manifest itself within a university laboratory setting, in which every effort was made to relax the participant. Despite the considerable thought and planning that went into Studies One and Two, they were not without a significant flaw - and ironically it is the success of the well-designed methodology that caused it. To eradicate sensory leakage and the possible artefacts that could account for positive results, the researcher created an experimental environment in which almost every variable could be controlled and manipulated. Whilst this was intended to ensure that the variables under investigation were affecting surveillance detection should it occur (rather than unintended influences), in doing so, the experimenter created an artificial situation which would never be encountered in everyday life. As people have reported sensing being stared at in real-world settings, the problem of ecological validity exists with the first two studies due to the laboratory setting which can only be overcome by conducting research in more realistic situations.

Through the use of deception and not telling the participants that they may be under surveillance in Study Two, the experiments were made more naturalistic - and it was under these conditions that the participants were more able to accurately self-report when they were being watched and listened to. Additionally, the element of being tested appears to have created a stressful environment more akin to the real world than the calm and protective setting of a laboratory, and it would seem that this stress caused participants EDA to increase in response to covert surveillance when there was no significant change if they were not being tested.

Study Three therefore investigated whether a more naturalistic setting could enhance the ability hinted at in the first two studies. This was also intended to potentially build on the results of Study One (see section 2.5.3) and Study Two (see sections 3.4.3 and 3.5.3) which demonstrated that task performance produced the most convincing data to indicate that scopaesthesia may exist - possibly as these were the only protocols which featured an element the threat, albeit mild.

The idea that threat may enhance individuals' remote surveillance detection ability led to the decision to include light levels as an additional variable in Study Three. It was deemed possible that participants would find the atmosphere created by lower light levels more intimidating, and this was supported by the literature. Farnworth et al. (2020) identified that bright light acts as an indirect visual cue to potential predation, which enables individuals to utilise correct risk avoidance behaviours to escape from threatening stimuli. Therefore, findings by Farnworth et al. (2020) indicate that bright light conditions offer much needed safety cues to detect danger in advance, whereas low light conditions may deprive individuals of safety cues needed to reduce the sense of threat.

From childhood, individuals develop a 'fear of the dark' (Gordon & King, 2002), with its origins suggested to emerge from a mix of different environmental factors (Rachman, 1977) such as vicarious learning (e.g. seeing an adult who displays fear of the dark) and information giving (i.e. repeated warnings from parents about dangers of being robbed or kidnapped at night). Gordon and King (2002) also argue that a 'fear of the dark' (or a sense of threat from low light conditions), is the result of biological preparedness, whereby individuals develop fears of potentially threatening stimuli from a very early age, which can also be increased through factors relating to attachment type such as separation anxiety (Schiele et al., 2020).

Other researchers have suggested that non-human animals develop predation risk schedules that are dictated by low light conditions (Tillman, 2009), whereby non-human animals exhibit a circadian shift in anti-predation behaviour, such as huddling together in groups and reducing movement at night in comparison to daytime activities. Therefore, threat generated by low light conditions or exhibiting a ‘fear of the dark’ seems to be a universal quality among both human and non-human animals, which suggests a sense of threat created by low light conditions has strong biological underpinnings such as biological preparedness to potential threats (Gordon & King, 2002).

With the current researcher having theorised based on the results of Study Two (see section 3.4.3) that ‘threat’ may be a necessary component for scopaesthesia to take place - light levels appeared to be a reasonable and ethically acceptable replacement for genuine threat. The literature would appear to support this idea, and so it was measured and included as a possible predictor variable for the field research.

Few field experiments have been conducted to research behavioural reactions to remote observation, with the notable exception of Sheldrake’s (2003) study (see section 1.5.2) in which observers were hidden from the observed participants by a one-way mirror. This method produced a highly significant difference in the reactions of the participants between observation and non-observation. Sheldrake’s work appeared to suggest that if participants turning to face their observer represented a genuine psi effect, then the influence exerted by the human observers staring directly at the participants exceeded the effect of the video cameras’ constant electronic surveillance during the experimental period.

Sheldrakes’s study (2003) also supports the experimenter’s idea that a field study conducted in a real-world setting would be a more ecologically valid test of whether remote surveillance detection is possible than the previous laboratory-based research. As such, to test whether environments may play a role in scopaesthesia, field studies were conducted in two areas which were judged by a focus group to be distinctly different to each other in terms of atmosphere. This was intended to investigate whether a ‘safe’ environment produces a difference in scopaesthesic ability when compared to a ‘threatening’ environment.

4.1.1 Study Three Hypotheses.

Based on the results of Study One and Study Two, the following hypotheses were developed to be tested in a real-world setting.

Formal Hypotheses:

- Participants' covert surveillance self-reporting accuracy will be significantly greater than chance expectation within a real-world setting in each area.

- Overall self-reporting accuracy will be significantly greater in the 'threatening' area compared with the 'safe' area.

- Covert surveillance self-reporting accuracy will be significantly greater in each area when participants are asked to report surveillance, when compared to non-surveillance.

Exploratory Hypothesis:

- The ability to detect covert surveillance will be associated with the environmental and behavioural factors under investigation, as well as belief in and experience of anomalous phenomena, temporal lobe lability, and Schizotypy scores - although this is an exploratory hypothesis.

4.2 Methods

4.2.1 Design.

The current study utilised a multivariate mixed experimental design to investigate whether covert surveillance influenced participants' self-reporting accuracy and changes in their behaviour, and whether the setting (the area) and atmosphere (light levels) in which this takes place plays a role in their scopaesthesic and acoustathesic detection. Behavioural measurements including walking speed, clothing adjustments, pauses, and whether the participants under surveillance turned to look over their shoulder were taken during periods of randomised surveillance which participants could not be aware of via conventional means. Participants' ability to detect being watched and listened to was analysed for possible trends, patterns and relationships with the periods of actual surveillance - as were the variables under investigation which included anomalous experience, anomalous belief, levels of schizotypy, and their temporal lobe lability.

4.2.2 Participants.

There was no active recruitment due to the naturalistic design of this experiment. Instead, participants (all aged 18+) consisted of university students, employees and visitors who happened to be walking along the two selected areas of observation within a London University campus during the periods of experimentation, with consent being retrospectively gained as described in section 4.2.9. This produced 100 (see appendix YY) participants in total, all of whom produced useable data - however only some (just 22) followed through with providing their individual differences data. Participants consisted of 66 (66%) females and 34 (34%) males, ranging in age from 18 to 67 years (*M* = 35.87, *SD* = 13.08).

Within this sample, the participants could be broken down into two distinct groups; those who were walking within the 'safe' area, and those walking within the 'threatening' area. The sample walking within the 'safe' area consisted of 74 participants ranging in age from 18 to 67 (*M* = 36.78, *SD* = 13.66) and 45 (61%) were female and 29 (39%) were male (see appendix ZZ). The sample walking within the 'threatening' area consisted of 26 participants ranging in age from 18 to 63 (*M* = 33.27, *SD* = 11.08) and 21 (81%) were female and 5 (19%) were male (see appendix AAA).

4.2.3 Materials.

The study required:

i) A discrete and weatherproof video camera (a Go Pro 5 Session) to record participants' behaviour so that it could be independently judged.
ii) A directional microphone to listen to the participants.
iii) A programme (Qualtrics) to administer the questionnaire.
iv) Two suitable environments free from hazards and distractions in which the observation should take place (see Appendix BBB for the 'safe' location, and Appendix CCC for the 'threatening' location).
v) SPSS statistical analysis programme to analyse the data.
vi) Apple's Lux Light Meter app to measure the amount of light (tests determined that the app worked most reliably when the phone was laid flat and facing upwards).
vii) Stopwatches for the judges to time how long it took the participants to walk the designated distances.

4.2.4 Focus Group and Pilot Studies.

To ensure the experiments met with the ethical standards of the British Psychological Society (BPS) as well as the University of Greenwich where the study was taking place, a focus group was held with six group members. These were evenly divided between university students and staff - this proportion was deliberate to try and replicate the intended study sample as the experimenter predicted that due to the research being conducted entirely on university grounds it was very unlikely that members of the public would form part of the participant group. Whilst this does compromise the generalisability of the results, this was necessary to adhere to the guidelines insisted upon by the University ethics committee.

The focus group was essential due to the unusual methodology being employed - namely that participant consent would be sought *after* the fact, rather than in advance of the study taking place as would usually be the case. This unorthodox approach was unavoidable though if the integrity and validity of the experiment was to remain intact, as its inception was driven by the need to investigate whether the results of Studies One and Study Two could be replicated in a real-world setting. The second experiment built upon the first by *not* informing (and therefore priming) participants to self-report whether they were under surveillance - and this experimental amendment demonstrated significantly different results. The first two studies were both conducted in a laboratory, and this controlled environment allowed the experimenter to ensure that the participants could not be aware of surveillance via conventional means as a one-way mirror and headphones were used to prevent sensory cues.

Such experimental luxuries as controlling the environment and limiting participants' distraction were sacrificed in Study Three in exchange for the naturalistic environment which offers increased validity, and therefore the opportunity to prevent unnatural behaviour from the participants, such as searching their surroundings for indications of surveillance. These indicators could include cameras and possible observers, and to avoid priming the participants to look for such cues, they were only approached *after* the period of surveillance. It is for this reason that a focus group was necessary as the experimenter wanted to investigate two topics related to this; firstly, that participants' reaction to this element of the study was unlikely to cause concern, offence, anxiousness, or any other negative reaction to ensure the well-being of those being observed and recorded. Secondly, it was essential to the success of the experiment that potential participants would be likely to give their consent - as without it, the data gained would be unusable and would need to be destroyed.

Whilst without exception all members of the focus group indicated that they understood the ethical concerns of the researcher, they also understood the methodological reasoning for consent being gained after the possible period of surveillance and had no issues with the experiment being conducted in this way. Interestingly, half (three) of the individuals in the focus group assumed that they would be being recorded in a public space anyway. All focus group participants were also asked;

a) If they were a genuine participant, would they give their consent for the recordings (visual and audio to be used). All focus group participants indicated that they would.

b) If they were a genuine participant, would they be willing to take the time to complete the rest of the experiment by completing the online demographic and individual differences questionnaire. All but one of the focus group participants indicated that they would.

c) If they were a genuine participant, would they consider £5 remuneration an adequate incentive to spend approximately 15 minutes completing the online demographic and individual differences questionnaire. All of the focus group participants indicated that they would.

d) Whether they considered the two areas in which the study would be conducted to be notably different from one another in terms of atmosphere, and that one was distinctly more threatening than the other.

At the end of the focus group session, those taking part were given the opportunity to voice any concerns they had with any other elements of the research as the experimental protocol was described to them in detail. Aware that some group members may be uncomfortable raising issues directly or in front of their peers, they were also given the opportunity to email the experimenter within seven days. No such emails were received. The focus group was concluded with the experimenter thanking the individuals who took part for their time. Satisfied that their hypothetical reactions to the conditions which would be experienced by genuine participants would be either positive or neutral, the experimenter proceeded to organise a pilot study to address possible physical complications.

4.2.5 Pilot Study One (Camera Positioning and Focus).

Expecting there to be methodological issues and complications due to the complexity associated with field studies, the experimenter took the role of both the researcher and the participant to enable identification of preliminary issues to ensure the efficient and smooth running of the experiment. The first problem encountered was the positioning of the camera intended to record the participants as they travelled between the start and finish points in each of the two locations. To keep the camera adequately hidden so that it could not be seen by passers-by who may then behave differently in its presence (and in doing so may alter the behaviour of the participant), the camera had to be some distance away from the area of surveillance. This was essential as if conventional awareness of the camera was gained, this could account for the very differences in behaviour that the experimenter was investigating (rather than the unconventional detection of surveillance as intended). Possible issues arose when having filmed the area upon which the camera would be trained from an adequate distance that it could not be spotted, the footage was from too far away to analyse it in detail. This methodological problem was overcome however via the employment of a videographer who could then zoom in on the area of interest and edit out the extraneous detail, therefore making the behaviour of individuals filmed within the space easier to examine.

4.2.6 Pilot Study Two (Identifying Suitable Areas).

Prior to the experiment taking place, suitable environments and vantage points needed to be explored and decided upon by the experimenter. Following consultation with the university ethics committee, it was determined that these environments should be within university grounds; this was so that when approached by the experimenter, any individuals under observation would be more likely to accept that the events taking place were indeed part of university-based research. Such safeguards were intended to reassure potential participants that the study was genuine, but also to protect the experimenter from possible aggressive reactions if any individuals felt threatened or suspicious of the activity.

The specific areas in which to conduct the field study were carefully chosen to meet various important criteria. They had to be hazard free so that the researcher could rest assured that any distraction which may be caused by the remote surveillance would not inadvertently lead to an accident. The area also had to be busy enough in terms of human traffic so as to provide enough participants for the study's efficiency, yet not to the extent that it was constantly travelled by groups of people who may cause a distraction and therefore interfere with the potential detection of observation. Whilst these areas needed to be outside (so that differing light levels could be taken), they also had to be within the proximity of an indoor area with a window through which the observer could stare at and listen to (during the surveillance conditions) passers-by who would (with their permission) become participants. This same window would also allow the video recorder to capture the participants' behaviour for subsequent analysis.

The two locations in which the research was to be conducted were selected for their very different qualities. To test for whether environment and atmosphere may play a role in participants' ability to detect surveillance, university grounds were explored for areas which differed from each other as much as possible in an effort to define one as a 'safe' location and the other as a 'threatening' location. This was due to the experimenter theorising that that the 'need' to detect surveillance predicts individuals' ability to do so - a theory supported by the results produced by the methodological redesign in Experiment Two (see section 3.4.3) when participants' physiology appeared to change when under surveillance, but only when under the pressure of being tested.

The 'safe' location was eventually decided upon due to its proximity between major and well-used buildings within the university, meaning that it was well travelled and familiar to most individuals who use it. Being a light, bright and open location with space either side of the well-defined path meant that potential participants were very unlikely to be concerned about unknown hidden persons, and would probably be reassured by the comparative busyness of the surrounding buildings. This was in stark contrast to the 'threatening' location, which was far more remote, enclosed, and seldom used which created a more foreboding atmosphere where the experimenter predicted a greater level of threat may be felt by those walking through.

The experimenter's assumptions were tested by the focus group, as well as 12 university students who were taken to both locations and asked hypothetical questions related to the apparent safety of both environments. The focus group concurred with the experimenter's opinion, and the vast majority of the students (83%) suggested that there was a notable difference between the two locations in relation to their atmosphere. The experimenter therefore decided upon these two areas due to the previously mentioned atmospheric differences between them, but also as they both featured a straight and seemingly equal length of path that would allow direct comparisons as they guide potential participants in a clear straight line.

To further test these areas for compatibility and for possible methodological issues, the experimenter enlisted the assistance of research students within the university to play the role of the participant. This offered the researcher the chance to physically test the experiment, and this revealed another possible methodological flaw; there was a substantial difference in the time it took the pseudo-participants to walk the pathways that would be under surveillance. This issue had not previously been salient, as the differences were somewhat hidden by random people with various walking speeds travelling along the paths. However, when the *same* individuals were traversing these areas, it became apparent that the length of the paths may not be as comparable as they first appeared. Measuring both paths demonstrated this to be true.

The seemingly obvious answer was to only analyse the first section of the longer path, so that the areas of surveillance were directly comparable; however, there was no obvious landmarks to indicate this distance to assist the independent judges who would later use the video footage to compare the participants' behaviour. To overcome this concern, the experimenter considered the use of physical markers to indicate exactly where the area of observation should begin and end. This idea needed to be re-thought however, as subtle markers such as sticks on the ground that would not stand out as suspicious to passers-by proved to be invisible when watched via the video footage. Conversely, anything that stood out on the video footage by way of being colourful or large enough to be identifiable from a distance was deemed to hint to participants that something unusual may be taking place - especially as the majority would be university staff and students who would know the spaces well, and so would be familiar with their layout.

An alternative way to make the two areas of surveillance comparable was found when via the use of a meter stick, the experimenter was able to establish that a clearly defined part of the path in both areas was of equal length. Tests showed that it took a person of average height walking at a normal pace (not pausing or rushing) approximately 30 seconds to get from the starting point, to the finishing point. This was deemed acceptable by the experimenter as if the periods of surveillance were too short; the participants would not have a reasonable amount of time to demonstrate behavioural change (if they were going to do so). For this reason, a greater length of time may have been preferable, but it was not possible due to the limits of the video camera's field of focus to record more than the 40-meter distance these comparable areas covered.

This comparable measurement also resolved another important theoretical methodological issue - namely that should scopathesia genuinely exist, then the research assistant's observation may be detected and elicit the very behavioural change intended to be generated from the experimenter's surveillance. The shortening of the surveillance area allowed the research assistant to remain out of the participants' view until the period of surveillance was over, to remove the possible artefact of the wrong person's surveillance being detected. This would also remove a possible 'doubling-up' of observation if both the experimenter and the research assistant were watching the participants, as this could affect the intensity of the stare - and in doing so could influence detection.

The doubling-up of observation also justified the use of the video recorder, as it was essential for various important reasons. The presence of the video recorder allowed the participants' behaviour to be analysed at a later date so that movements and gestures could be examined even when the experimenter was not staring and listening (and would otherwise have no way of knowing how the participants' behaviour compared with when they were being stared at and listened to). It is important to note that the use of video recording in this experiment was unavoidable, as by simply having an assistant monitor the participants behaviour as it occurs, the assistant would need to directly watch - and therefore would effectively be creating an 'observation' condition even when the experimenter is conducting a non-observation condition. Fortunately, concerns that the video camera itself may cause participants to feel watched were eased by Sheldrake's research (2003) which tested participants' behavioural reaction to being remotely observed, as this demonstrated that direct viewing was more detectable than CCTV when the cameras seemingly had little effect on individuals' ability to detect being stared at.

Just as importantly, the filming of the participants' behaviour allowed repeated viewing and for it to be independently judged by two other individuals (not the experimenter) for the behavioural changes (quickening of pace, the participant turning to face the observer, adjustments to clothing, and pauses). These behavioural changes needed to be coded (see section 4.2.8) so that the judges had a clear definition of exactly what constituted a 'change' in behaviour, as otherwise one judge may deem a participants' behaviour as meaningful, whilst the other may not. These scores could then be compared with each other to ensure a consensus was reached.

4.2.7 Pilot Study Three (Recording Data and Testing).

With the field experiment significantly closer to being ready, its methodology was again tested in a further pilot study that included all of the changes and improvements made up until that point - but also incorporated the complications associated with coordinating the experiment, as well as the analysis to be conducted on completion. It was deemed necessary for the research assistant to note participants' distinguishing features via codes (see appendix DDD) which could be generated quickly, but would be meaningful later and enable the experimenter to identify each participants' individual footage. This became necessary when the original idea to identify participants solely through the order in which they arrived at the location became redundant as the third pilot study revealed further matters to be addressed due to distractions, and the unprecedented number of people engaged in conversation. These conversations could be with people around the participant, or on a mobile phone - but in either case, it was decided that these individuals should not form part of the sample. This was to ensure closer adherence to the experimental protocol of studies one and two, as well as the majority of the studies detailed in the literature (in which participants were alone).

Interestingly, the changing weather conditions that occurred whilst the Study Three methodology was being tested helped to eliminate artefacts causes by unintended environmental conditions, as it was noticeable how people's pace quickened generally when it rained - for this reason it was decided that the experiment would only be conducted in dry conditions. Whilst theoretical, it was also decided that the experiment would not be conducted in icy conditions as this would be likely to cause a slowing of pace that would not be due to the variables under investigation. As such, the experimenter determined that to qualify, participants must:

a) Be alone (other individuals can be on the same pathway, but not talking to the participant or part of the same group).

b) Not be on a mobile phone.

c) Not be listening to headphones.

d) Not be distracted by a notable disturbance in the vicinity (as determined by the research assistant) to be significant enough to draw the participants' attention to the extent that possible extrasensory detection may be over-ridden.

e) Be walking in dry, ice-free conditions.

The research design was finally tested via a pilot group (N = 12) who experienced the same procedure and conditions intended for the actual participants to ensure that the experimental aims were not obvious whilst it was taking place, and that the observer could not have been detected via conventional senses. None were able to guess the experimental aims before it was explained to them, and any opportunities participants may have had to gain awareness of being watched via conventional senses were identified and subsequently eliminated. This pilot group was also intended to further investigate potential discomfort, but no concerns or issues were reported.

4.2.8 Coding the Behavioural Changes.

With the task performance protocol results from Study One and Study Two both indicating that behavioural change may not only be a means of detecting covert surveillance, but in fact may be a more reliable measurement than self-reports - this variable was included in the field study. The experimenter wanted to stand the best possible chance of capturing the phenomenon of remote surveillance detection should it exist, and so included both self-reports and behavioural change as means of possible capture. This would also allow for cross-referencing of potential patterns in the data; for example, if there were significant changes to self-reporting accuracy *and* participants' behaviour in a particular condition - this could be considered even more important than if positive results were found by just one measure. For this reason, the interpretation of participants' behavioural changes could not left to the discretion of the experimenter or the research assistant due to possible bias, the fallibility of memory, or lack of attention whilst they attended to the running of the experiment. This is one of the reasons the experiment was recorded.

Recording the participants' allowed repeated viewing, and for any behavioural changes (differences in walking speed, the participant turning to face the observer, adjustments to clothing, and pauses) to be independently judged by two individuals not otherwise involved in the study. In anticipation of possible differences in what the judges deemed meaningful; clear and simple definitions of exactly what constituted a ‘change’ in behaviour were established. These were as follows:

- Walking Speed - this was judged by the number of seconds it took for participants to walk from the starting point, to the finishing point. These positions were highlighted on a video screen by the experimenter to both judges to avoid confusion using a test participant who was not included in the study to allow the judges a chance to practice their timing using the stopwatches provided. The times were rounded up or down to the closest whole second to allow a realistic chance of the judges’ timings matching. It was considered that a quickening of pace would indicate the participant may be detecting observation, as this was assumed to indicate the unease often associated with the feeling of being watched (Coover, 1913; Sheldrake, 2003a; Thalbourne & Evans, 1992; Titchener, 1898).

- Head Turns - these were judged by any turning of the participant’s head of between 45 and 180 degrees, as the observation was positioned from behind. If this occurred, it was recorded by the judges. A head turn of less than 45 degrees was deemed to not be a response to a feeling of being under surveillance, as this would not have allowed the participant to look in the direction from which the stare was emanating. For the purposes of the study, it was considered that a registered head turn indicated a response to covert surveillance.

- Clothing adjustments - these were included as an indicator of self-consciousness, which may be experienced if one felt as though they were being watched. As such, all notable adjustments to the participants’ clothing were included as possible indicators of covert surveillance detection and counted. Although the experimenter notes that factors including the temperature, the type of clothing, and whether there was a breeze could all influence this - an adjustment of participants’ clothing was considered to mean that remote surveillance had been detected.

- Pauses - the participants stopping for no obvious reason was thought to be a possible indicator of covert surveillance detection. The amount of time each participant did this was counted, with a higher score considered to mean an increased feeling that surveillance was taking place. To be judged as a pause, the participant needed to come to a complete stop, for any amount of time. Overly long pauses would have been noted as they would have affected the 'walking speed' variable to the extent this may have created outliers - however this never happened. If the pause was for an obvious non-surveillance related reason this was to be discounted - however, in reality, the research assistant never included trials with major events or disturbances that caused noticeable distractions - and so this potential problem was averted.

It is important to note that the experimenter did not view the recorded footage until after the judges had completed their task and submitted their scores, so that they could not be influenced to meet expectations or experimental aims. The details of the task were explained to the judges (both final year students at the university where the study was conducted) so that they received identical instructions, however, they conducted their judging sessions in separate rooms to ensure there was no harmonising of their decisions - either intentional, or accidental. Similarly, the experimenter was not present during the judging sessions to ensure that outcomes could not be inadvertently influenced. Only after both judges had completed their tasks did the experimenter return to examine whether there were discrepancies between the timings and the behavioural scores given. Should there have been no consensus, a third independent judge would have been recruited - however, it seemed the clarity of the coding instructions had the intended effect, as both judges were in 100% agreement.

4.2.9 Procedure.

Due to the complex ethical considerations necessary for opportunistic observational based research, not only were all the standard British Psychological Society (BPS) Code of Ethics and Conduct of briefing participants, inviting and answering their questions, obtaining consent, ensuring confidentially, informing them that that they can withdraw themselves and their data at any point, and debriefing them fully adhered to in retrospect, but further precautions were taken. To increase confidence that the study would not cause distress, a focus group was held (see section 4.2.4 for details) to uncover potential issues. This was intended to inform the protocol and minimise participant discomfort, however no such concerns were raised.

To qualify to become a participant, an individual had to be walking forwards (from the observers' perspective) within the designated observation area, and to have spent at least 30 - 60 seconds doing so, as during pilot studies the experimenter estimated this to be a) a reasonable amount during which measurable behaviour may occur, and b) the amount of time it would take most people to travel the 40 metre path from start to finish in both locations.

To not add another layer of difficulty to an already complicated technical process, surveillance and non-surveillance trials were not randomised as this would have been yet another element to keep track of during a constantly evolving situation - however, this was not a detriment to the quality of the study. As there was no contact or communication between the experimenter and the participants, it was impossible for the observer to accidentally or deliberately prime them as to which condition to expect, as the participants were unaware of the experiment until after it was completed. Importantly though, the research assistant was unaware of when the surveillance trails were taking place, so that this could not be conveyed to the participant through word choice, tone, or body language - as it was the research assistant who asked the participant whether they felt like they had been watched or listened to. Because the participant recruitment was opportunistic, the experimenter was not able to allocate them to groups as there was no control over who walked into the experimental area. This meant that groups were randomly allocated - albeit by chance. They could not guess surveillance patterns as there were no patterns to guess, and they were not aware of the need to do so until the opportunity had passed.

The observers' vantage point at the 'safe' location was chosen mainly because it allowed a clear view of the passers-by, but only once they reached the position at which monitoring begun and they effectively became participants. Until this time, the observer could not see the participants - however these individuals were in view of the video camera. This allowed the judges to examine the behaviour of the participants prior to them entering the area of observation so that behavioural comparisons could be made with when they come into the observers' view.

The vantage point was also important so that during the observation conditions, the staring took place from behind the participants to insure they could not become aware that they were being watched via conventional means (see figure 5). This positioning was then also synonymous with the majority of the literature - in particular, the classic studies in which one participant stares at the back of the other's head or neck (Coover, 1913; Lobach & Bierman, 2004; Poortman, 1959; Schwartz & Russek, 1999; Sheldrake, 1998, 1999, 2000, 2001; Titchener, 1898).

Once the passer-by had reached the point at which monitoring ended, a research assistant approached them and explained the experiment during which time he/she informed them that they had been filmed, and possibly observed. The participants were then asked if they consented to being part of the research; if they declined, they were thanked for their time, the session was terminated, and the recording was permanently deleted. If they said 'yes' however, the light level was recorded (so that this could later be analysed as a variable), and they were invited to complete a consent form before being given information by the research assistant which included a link to the demographic and individual differences questionnaire (created in the Qualtrics software programme. It was hoped that the individuals would then complete this questionnaire (estimated to take 15 minutes) at a computer or via their phone, and an incentive of £5 remuneration for their time and effort was offered to increase the chances of them doing so, which could be collected upon completion. As the researcher predicted that not all participants who consented to take part would complete the follow-up to the study - to allow for the important self-report measure to be guaranteed, participants were also asked if they 'felt' as though they were being watched and listened to. By having the research assistant gain this valuable information in person, the self-report responses were assured even if the participant failed to complete the questionnaires later.

Once the participants who consented to take part were thanked for their contribution, the debrief form was then issued and participants were reminded of their freedom to withdraw their data, and that they should seek support if they found any elements of the study unsettling. Whilst the importance of not relaying details of the experiment to anyone else was specifically mentioned in the debrief form, this was also stressed verbally before the participant left as should the word spread regarding the experiment and subsequent passers-by consequently become aware of the experimental aims it could jeopardise the intended naturalistic methodology. Any individuals who decided against being included in the research were instantly thanked for their time and left to continue on their way with the assurance that the footage related to them would a) not be used as data, and would b) be destroyed. In either instance, the research assistant noted enough distinguishing features to identify the individual in the video so that their footage could either be analysed in conjunction with their demographic details and self-reported awareness of observation, or deleted as requested.

Consenting participants' self-reported awareness of surveillance and the differences in their behaviour was subsequently compared against whether and when they actually *were* under observation to investigate possible relationships and patterns with the variables under investigation via and t-tests. The below diagram is intended to make the experimental set-up clearer.

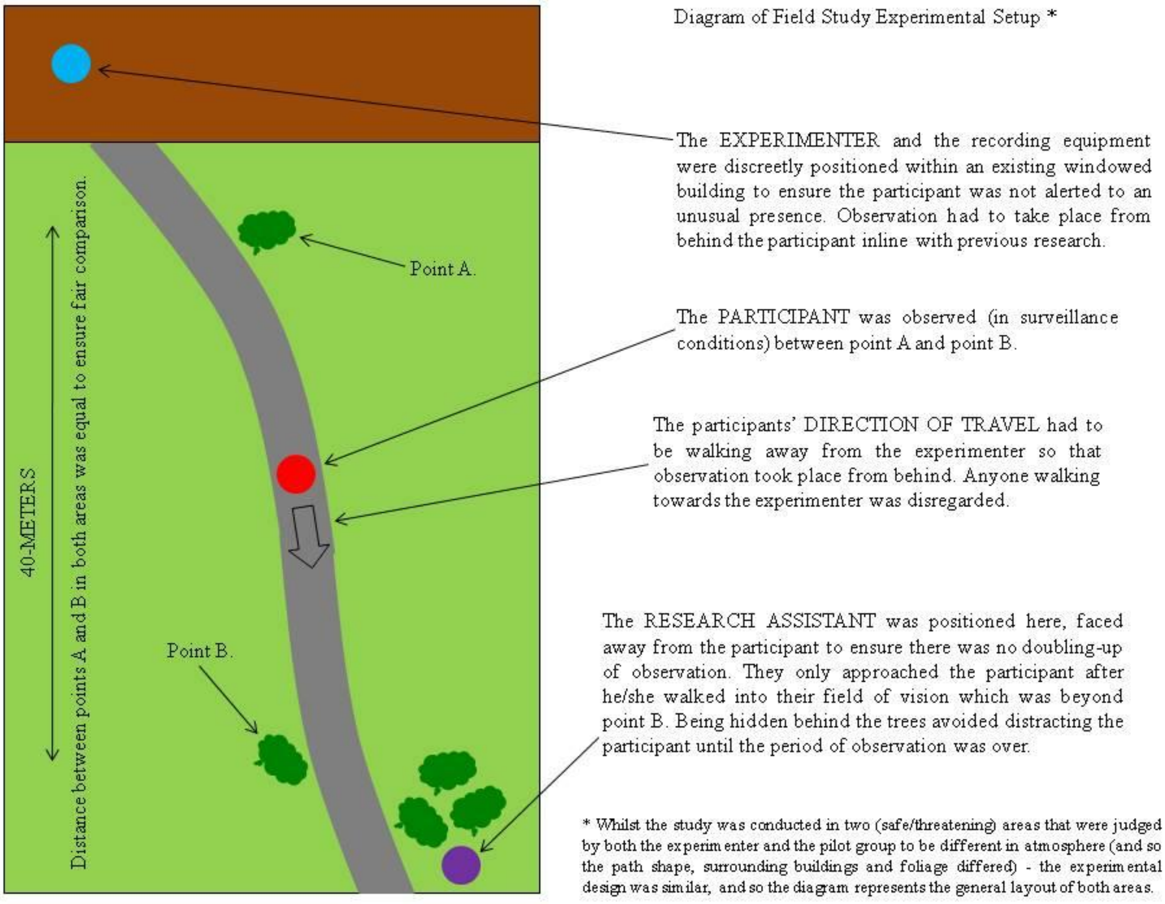


Figure 5 A diagram to show the experimental layout of Study Three.

4.3 STUDY THREE RESULTS

4.3.1 Analysing the Self-Report Scores for Possible Bias.

There were no missing data points within the results, and table 25 below shows descriptive statistics for the self-reported feelings of surveillance (see appendixes EEE, FFF and GGG) of Experiment Three for all conditions. It is important to note that the below figures do not denote correct responses, but instead represent the amount of times participants said they felt they were under surveillance. In reality, overall participants were under surveillance half (50%) of the time.

Interestingly, 53% of participants reported feeling as though they were under surveillance overall which is not statistically above chance expectation (p = .276), whilst 51.4% of the participants in the 'safe' condition reported the same sensation - which again was not statistically above chance (p = .409). However, a greater percentage (57.7%) of the participants in the 'threatening' condition reported the feeling of surveillance, but due to the lower number of participants (N = 26) this was also not significantly above chance expectation (p = .222), so there was not a significant reporting bias in this third experiment.

Table 25

Descriptive Statistics of How Often Participants Reported Surveillance

Self-report conditions	Mean	Std. Deviation
Safe Area (N = 74)	0.51	0.503
Threatening Area (N = 26)	0.58	0.504
Overall (N = 100)	0.53	0.502

Note. The above does not indicate correct reports, but rather the amount of times surveillance was reported.

4.3.2 Comparing Correct Self-Reports With Chance Expectation.

The data was analysed to investigate the formal hypothesis that participants' self-reporting of covert surveillance accuracy will be significantly greater than chance expectation in a real-world setting in each area. The below tables show the descriptive statistics of participants correct responses for both conditions (surveillance and non-surveillance) in each of the designated areas (safe and threatening) of the field study (see appendixes HHH and III), as well as the correct self-reports of both conditions combined. These tables also demonstrate how the correct self-reports means compare with chance expectation. It is important to note that chance expectation was 25% for 'correct reports of surveillance' and 'correct reports of non-surveillance' as the mean represents how many times participants were correct in reporting *either* surveillance or non-surveillance only. The 'correct reports overall' mean represents the total number of participants' correct self-reports for surveillance *and* non-surveillance combined, hence chance expectation being 50%.

Table 26

Correct Self-Report Descriptive Statistics in the Safe Area

Correct Self-Report Type	Mean	Std. Deviation
Correct Reports Overall	.50 (chance = .50)	.503
Correct Reports of Non-Surveillance	.24 (chance = .25)	.432
Correct Reports of Surveillance	.26 (chance = .25)	.440

Table 27

Correct Self-Report Descriptive Statistics in the Threatening Area

Correct Self-Report Type	Mean	Std. Deviation
Correct Reports Overall	.69 (chance = .50)	.471
Correct Reports of Non-Surveillance	.31 (chance = .25)	.471
Correct Reports of Surveillance	.38 (chance = .25)	.496

4.3.3 Analysing the Self-Reports for Comparisons With Chance Expectation.

As the data did not pass the assumptions for one-sample t-tests, the self-report scores for the 'safe' and the 'threatening' areas were examined for significant differences with chance expectation using binomial tests (see Appendix JJJ and Appendix KKK). As previous research detailed in the literature had not made this environmental distinction, these tests were intended to see if the self-report results generated within laboratory conditions could be replicated in a real-world setting.

SAFE AREA

A binomial test was conducted to compare participants' overall correct self-report scores against chance expectation whilst in the 'safe' area. The difference was non-significant as the test indicated that the proportion of correct self-reports of .50 was not above chance expectation .50, $p = .50$ (one-tailed).

THREATENING AREA

A binomial test was conducted to compare participants' overall correct self-report scores against chance expectation whilst in the 'threatening' area. The difference was significant as the test indicated that the proportion of correct self-reports of .69 was above chance expectation .50, $p = .039$ (one-tailed).

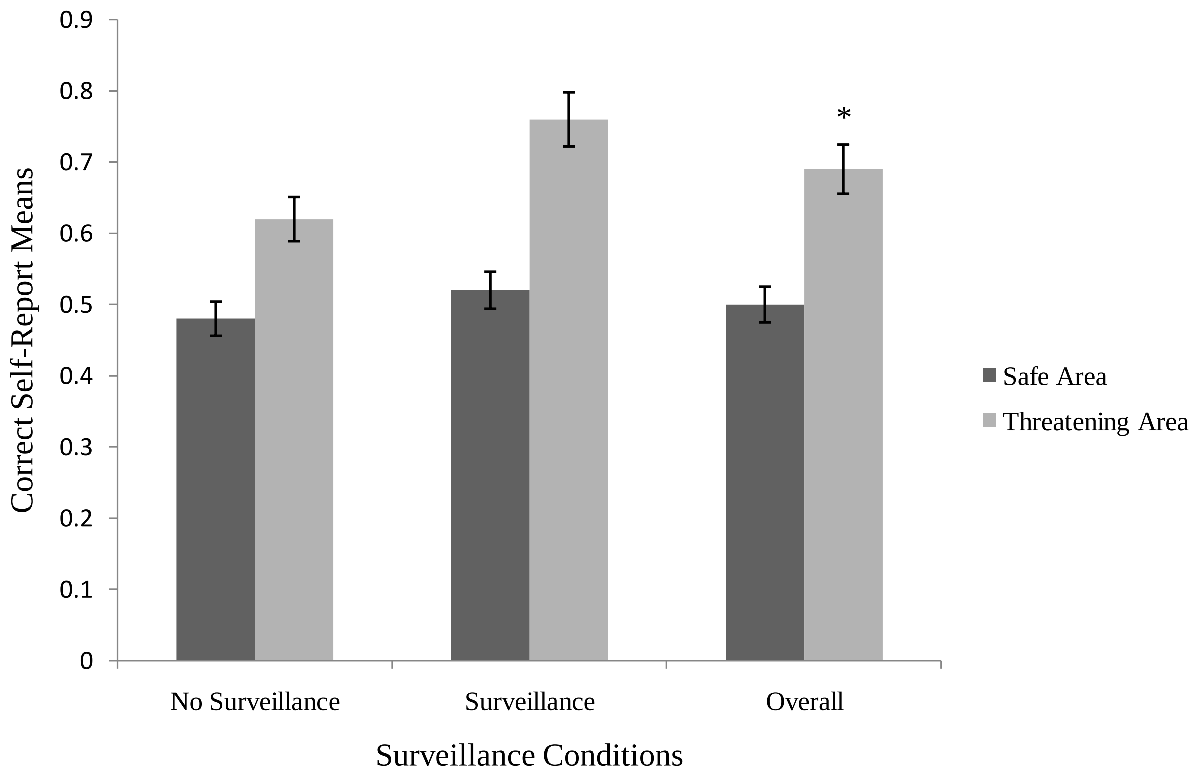


Figure 6 Self-report means compared to chance expectation.

Note. The 'No Surveillance' and 'Surveillance' means have been doubled to be visually comparable with 'Overall'.

*. One-sample t-test significant at the 0.05 level (one-tailed) when compared with chance expectation.

4.3.4 Analysing the Self-Reports to Compare Accuracy Between Areas.

To test the hypothesis that participants' overall self-reporting accuracy would be significantly greater in the 'threatening' area compared with the 'safe' area, a chi-square test was conducted to examine in which area participants self-reported covert surveillance most accurately. Below are the descriptive statistics followed by results of this chi-square test (see Appendix LLL).

Table 28

Field Study Self-Report Descriptive Statistics

		AREA		
		Safe	Threatening	Total
Correct	NO	37	8	45
	YES	37	18	55
	Total	74	26	100

Differences between overall correct self-reports in the 'safe' area and overall correct self-reports in the 'threatening' area. A chi-square test of independence was performed to compare the overall correct self-reports in the 'safe' area against the overall correct self-reports in the 'threatening' area. Participants were found to be significantly more likely to self-report correctly in the 'threatening' area than in the 'safe' area, X^2 (1, N = 100) = 2.875, p = .045 (one-tailed).

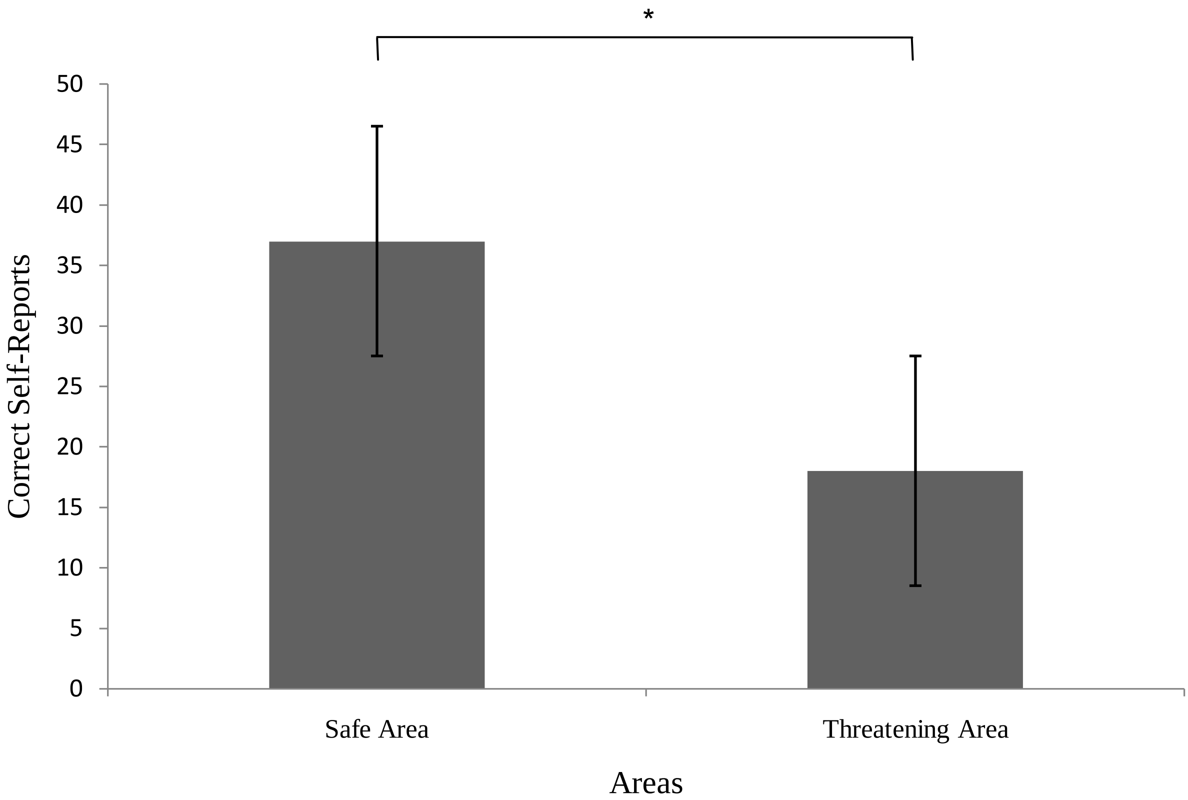


Figure 7 Differences in overall self-reporting accuracy in each area.

Note. Participants' self-reporting accuracy was higher in the 'threatening' area when compared to 'safe' area'.

*. Chi-square test significant at the 0.05 level (one-tailed) with participants found to be significantly more likely to self-report correctly in the 'threatening' area than the 'safe' area.

4.3.5 Analysing the Self-Reports to Compare Accuracy Within Areas.

To test the hypothesis that covert surveillance self-reporting accuracy would be significantly greater in each area when participants self-reported surveillance, compared to when they self-reported non-surveillance, chi-square tests were conducted for each of the areas (safe/threatening) to examine participants' self-reporting accuracy. Below are the descriptive statistics followed by the results (see Appendix MMM) of the analyses.

Table 29

Field Study Self-Report Descriptive Statistics for the 'Safe' Area

		Correct self-reports in the 'safe' area		
		No	Yes	Total
Actual Surveillance	NO	19	18	37
	YES	18	19	37
	Total	37	37	74

SAFE AREA *Correct responses of surveillance compared to correct responses of non-surveillance.* A chi-square test of independence was performed to compare correct self-reports of surveillance against correct self-reports of non-surveillance in the 'safe' area. Participants were not significantly more likely to self-report surveillance correctly than non-surveillance, X^2 (1, N = 74) = .054, p = .41 (one-tailed).

Table 30

Field Study Self-Report Descriptive Statistics for the 'Threatening' Area

		Correct self-reports in the 'threatening' area		
		No	Yes	Total
Actual Surveillance	NO	5	8	13
	YES	3	10	13
	Total	8	18	26

THREATENING AREA *Correct responses of surveillance compared to correct responses of non-surveillance.* A chi-square test of independence was performed to compare correct self-reports of surveillance against correct self-reports of non-surveillance in the 'threatening' area. Participants were not significantly more likely to self-report surveillance correctly than non-surveillance, X^2 (1, N = 26) = .722, p = .20 (one-tailed).

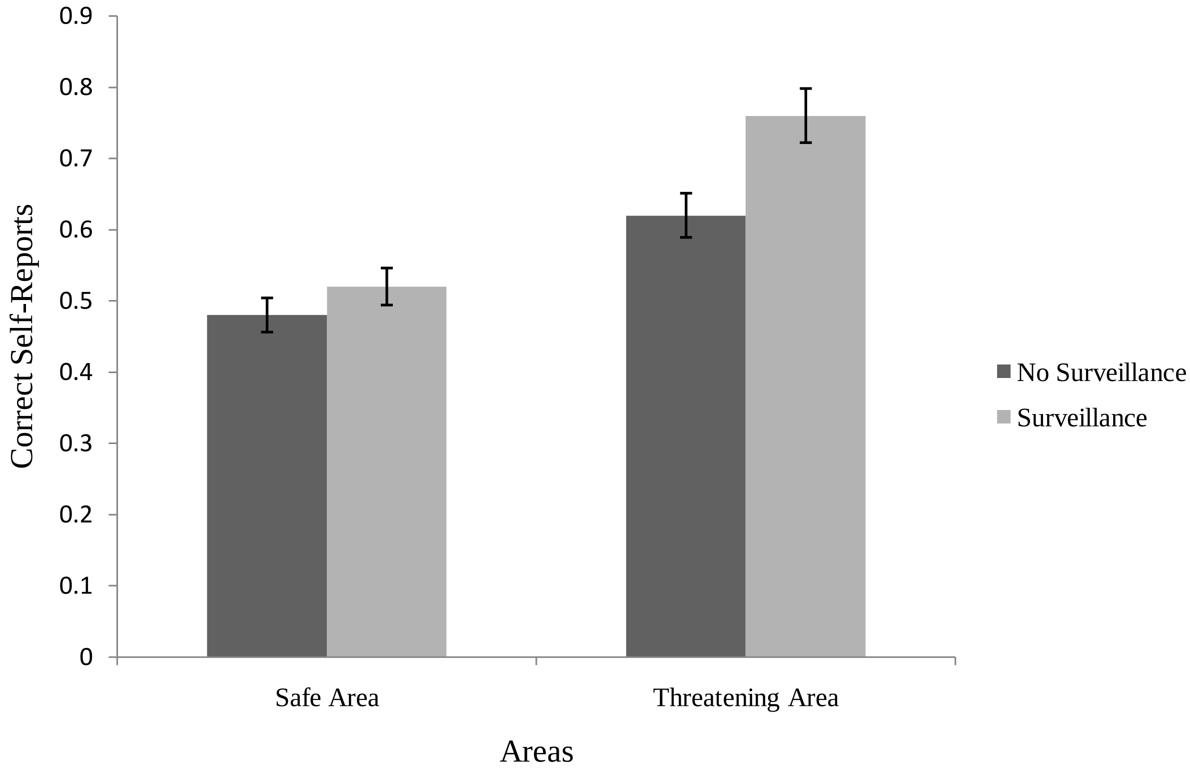


Figure 8 Correct self-report means compared to each other in both areas.

Note. Correct reports of surveillance and non-surveillance were not statistically different from each other in either area, but they were in the predicted direction with 'surveillance' yielding a higher number of correct reports than 'non-surveillance' in both areas.

4.3.6 Relationships Individual Differences, Behavioural and Environmental Variables.

Despite the apparent lack of relationships between covert surveillance detection and the psychosocial and neurological factors under investigation (belief in and experience of anomalous phenomena, temporal lobe liability, and schizotypy) demonstrated in the laboratory-based experiments of Study One and Study Two - the experimenter intended to include these individual differences variables as part of Study Three. These variables were not omitted, just in case the field study revealed them to play a larger role in covert surveillance detection if this was tested in a more ecologically valid setting; however, this element of the experiment was disappointing. Despite the effort the experimenter went to in running a focus group to ensure that participants would conclude the study in their own time by completing the online questionnaire that would capture their individual differences, and that the financial incentive offered would ensure this - in reality this was not the case. Only 22 of the 100 participants completed this element of the study, and of those, only 6 were tested in the 'threatening' area. This amount of respondents is too few for any meaningful analysis (Tabachnick & Fidell, 2007), and so the experimenter was forced to reluctantly omit these factors for the results.

Interesting individual differences measures remained however in the form of the light levels and the behavioural change factors - as these took place during the period of observation and so did not require further interaction on the part of the participants. With participants' behavioural reactions having been demonstrated to be an indicator covert surveillance detection via unconventional means in Study One and Study Two - it was hypothesised that they would have a relationship with self-reporting accuracy in Study Three. As there could be no cognitive task due to the covert nature of this final study, as well as the light levels taken during the experiment - the recording of the participants was carefully analysed for differences in their behaviour.

The pace at which the participants walked, whether they turned their heads towards the observer, whether they adjusted their clothing (assumed to be a sign of self-consciousness), and whether they paused whilst walking within the designated areas were analysed for relationships with self-reporting accuracy. The tables below show the descriptive statistics (see Appendix NNN) for these factors.

Table 31

Descriptive Statistics of the Behavioural and Environmental Variables

Predictors Measured	Correct Self-Report Mean	Incorrect Self-Report Mean	Correct Self-Report SD	Incorrect Self-Report SD
Light (in Lux)	12000	11000	1154	1425
Walking Speed (in secs)	30.10	29.44	3.510	4.885
Head Turns	.80	.31	.919	.602
Clothing Adjustments	.40	.63	.966	.806

4.3.7 Analysing Behavioural and Environmental Differences.

As the 'threatening' area was the only area that produced any significant results - this was the only area the experimenter examined for possible relationships with light levels and the behavioural measures under investigation. The independent judges agreed on the number and durations of the behavioural variables (see 4.2.8 for how these were defined and coded), however there were issues with the environmental and behavioural factors meaning the assumptions required for t-tests were not met (see appendix OOO).

The data for both the ‘Light Level’ and ‘Walking Speed’ variables were close to passing the assumptions of parametric tests, and were almost normally distributed despite a slight positive skew in the ‘Light Level’ data, and slight kurtosis in the ‘Walking Speed’ data. With t-tests considered to be robust enough to be valid for such distribution (Lumley et al., 2002), the experimenter decided that independent samples t-tests could be used. The data for ‘Head Turns’ and ‘Clothing Adjustments’ was non-parametric due to the limited score range (0 - 3) however, and so independent samples t-tests were not appropriate. A Mann-Whitney test could also not be used as there was not enough range in the data, and so these variables were analysed via chi square tests. Perhaps surprisingly, not a single participant slowed their pace to the point of stopping whilst walking in either the ‘safe’ or the ‘threatening’ areas and so there was no data with which to run an analysis for the ‘Pauses’ variable.

4.3.8 Analysing the Environmental and Behavioural Differences via T-tests.

Independent samples t-tests (see Appendix PPP) were conducted to investigate whether the light levels and the behavioural factors under investigation shared a relationship with reporting accuracy in the ‘threatening area’.

Light Levels. An independent samples t-test was conducted to investigate whether the light levels when participants were correct (M = 12000, SD = 1154) were different from the light levels when participants were incorrect (M = 11000, SD = 1425) in the ‘threatening’ area. The difference in light levels was non-significant between participants correctly self-reporting whether they were being observed or not, compared to those incorrectly self-reporting whether they were being observed or not; $t(24) = -.625$, $p = 0.269$ (one-tailed), d = .0.17 (-0.38, 0.71).

Walking Speed. An independent samples t-test was conducted to investigate whether participants’ walking speed when they self-reported correctly (M = 30.10, SD = 3.51) was different to their walking speed when they self-reported incorrectly (M = 29.44, SD = 4.89) in the ‘threatening’ area. The difference in participants’ walking speed was significant between those correctly self-reporting whether they were being observed or not, compared to those incorrectly self-reporting whether they were being observed or not, as the participants who were correct tended to walk slower than those who were incorrect; $t(24) = -1.907$, $p = 0.035$ (one-tailed), $d = .31$ (-.23, .86).

4.3.9 Analysing the Behavioural Differences via Chi-Square Tests.

Chi-square tests were conducted to investigate whether participants 'head turns' and 'clothing adjustments' shared a relationship with reporting accuracy. The descriptive statistics and test results can be seen below (see Appendix QQQ).

Table 32

Field Study 'Head Turn' Descriptive Statistics

		Turned to face the observer?		
		NO	YES	Total
Correct	NO	7	1	8
	YES	9	9	18
	Total	16	10	26

Head Turns. A chi-square test of independence was performed to examine the relationship between overall correct reports of surveillance and whether participants turned their heads toward the observer. The relationship between these variables was significant, with participants more likely to turn their head when they self-reported correctly than when they self-reported incorrectly, X^2 (1, $N = 26$) = 3.291, $p = .035$ (one-tailed).

Table 33

Field Study 'Clothing Adjustment' Descriptive Statistics

		Adjusted clothing?		
		NO	YES	Total
Correct	NO	4	4	8
	YES	13	5	18
	Total	17	9	26

Clothing Adjustments. A chi-square test of independence was performed to examine the relationship between participants overall correct self-reports of surveillance and whether they adjusted their clothing. The relationship between these variables was non-significant, with participants no more likely to adjust their clothing whilst under surveillance than when they were not, X^2 (1, N = 26) = 1.208, p = .136 (one-tailed).

4.4 DISCUSSION OF STUDY THREE

To investigate the formal hypothesis that participants' covert surveillance self-reporting accuracy would be significantly greater than chance expectation within a real-world setting in each area, the data was first checked for reporting bias. The participants were not primed in any way as they were not even aware of potentially being part of an experiment until after the fact, and so reporting bias was unlikely. Analysis demonstrated this to be true. At the end of their period of surveillance (or non-surveillance) participants were asked whether they felt as though they had been watched and listened to, and as they were choosing between two options, their correct responses were compared with chance expectation. The results showed that participants' self-reports in the 'safe' area did not significantly differ from chance expectation (being correct 50% of the time). These results in isolation would suggest that in a real-world setting that is not threatening, individuals are no more likely to be correct in detecting surveillance via extrasensory means than if they were just guessing randomly.

When participants' correct self-reports were compared with chance expectation in the 'threatening' area however, there was a statistically significant difference. Whilst these results only partially support the hypotheses, the experimenter considered these findings important as the threatening area is the *only* area in which the accuracy of participants' self-reports differed significantly from chance expectation. Whilst tenuous due to the limited number of participants attributed to this area (N = 26), this is nevertheless an interesting result as it supports the author's theoretical position that peoples' extrasensory ability to detect covert surveillance is dependent on circumstances, and that stress or threat is a necessary aspect of these situations.

To further investigate whether participants were able to detect covert surveillance in a real-life setting, the hypothesis that overall self-reporting accuracy would be significantly greater in the 'threatening' area compared with the 'safe' area was also tested. Participants' correct self-reports in each area were compared against one another for significant differences, and it was found that the hypothesis was supported as participants were more likely to self-report correctly in the threatening area. Whilst the relatively low number of participants in the 'threatening' area is less than ideal - this result suggests that the researcher may have been correct to choose to conduct the field study in two areas which featured distinctly different atmospheres, and to ensure that these atmospheres differed from one another enough to engage participants' extrasensory ability to a greater extent in one area than the other.

With two of the three formal hypotheses at least partially supported, the experimenter focused on the third formal hypothesis related to Study Three; namely that covert surveillance self-reporting accuracy will be significantly greater in each area when participants are asked to report surveillance, when compared to non-surveillance. Comparing correct responses of surveillance against correct responses of non-surveillance yielded non-significant differences in both the 'safe' and the 'threatening' areas meaning the hypothesis was not supported. It is worth mentioning though that analysis of the 'safe' area revealed the results to be no better than chance, whilst self-reporting accuracy was higher in the surveillance condition (77%) compared to the non-surveillance condition (61.5%) in the threatening area. It is therefore likely that this non-significant finding was most likely due to the small sample size (N = 26).

With interesting results related to self-reporting accuracy found, the experimenter investigated the exploratory hypothesis that participants' ability to detect covert surveillance will be determined by the environmental and behavioural factors under investigation, as well as belief in and experience of anomalous phenomena, temporal lobe lability, and Schizotypy scores. This prediction initially lead to disappointment, as despite the efforts the experimenter went to in ensuring participants would provide the required data by holding focus groups - only around a quarter did so. The likelihood of participants taking the necessary online questionnaire was investigated during the focus group session, and the vast majority (five out of six) concurred that they believed the financial incentive on offer would be sufficient to ensure most individuals would complete this element of the study. Unfortunately, this was not the case as only 26 participants of the total 100 took the time to conclude the study fully. Furthermore, with the analyses of the formal hypotheses showing that the safe area only produced non-significant results - to avoid multiple analyses, the experimenter intended to only investigate the exploratory hypotheses in relation to the threatening area. When looking more closely at the data from the field study revealed that of the 26 participants who undertook the questionnaire, only 6 were involved with the 'threatening' area data, the experimenter reluctantly accepted that meaningful analysis of the results would not be possible as the sample size was too small (Tabachnick & Fidell, 2007). Plans to examine the neurological and psychosocial individual differences were subsequently abandoned, and the experimenters' attention turned to the environmental and behavioural factors under investigation.

With participants' behavioural reactions having been demonstrated to be a viable indicator of individuals' ability to detect surveillance via unconventional means in Study One and Study Two, the experimenter was keen to investigate whether behaviour would have a relationship with covert surveillance of the participants in a natural environment. It was predicted that participants under surveillance would increase their walking speed, turn their head towards the observer, adjust their clothing, and that they may pause more often. These changes in behaviour were assumed to be signs of self-consciousness or nervousness which could be exhibited in response to the "uncanny… unpleasant tension" Titchener (1898, p895) described, and the restlessness, discomfort, feeling of being criticised and wanting to turn around that Coover (1913) reported to be feelings associated with scopaesthesia.

Either not all of the behaviours assumed to be indicators of covert surveillance were valid measures, or the methodology was not correct for these actions to manifest as analysis of these factors produced only two significant results. In fact, *none* of the participants paused during the experiment period at all, resulting in no data to analyse. Whilst participants' walking speed initially looked as though it may provide interesting data when some individuals took almost twice as long as others to walk the same distance, an independent samples t-test showed that whilst this behavioural variable was significantly linked with self-reporting accuracy - it was in the opposite direction with participants who walked slower tending to be more accurate with their self-reports. The experimenter noted though that participants increasing their walking speed in response to feeling as though they were under surveillance was an assumption, and was not based on research. Consequently, it could be that the feeling of being the subject of another individuals' focus actually makes it more likely that participants will slow their pace - possibly to be more cautious and hesitant when observed. Research aiming to answer this question would be interesting in its own right, but would also help with any future extensions or replications of this study.

Whilst previous to the research taking place, the experimenter wondered if any of the participants would adjust their clothes during the relatively short 40-meter journey that formed the observation area - a surprising 35% (9 out of 26 did so), with one individual doing so as many as 3 times. This behaviour appeared to be unrelated to reporting accuracy though, as whilst participants were *less* lightly to adjust their clothes when they were correct in their self-report - this difference was not significant. This non-significant relationship could of course simply mean that there is no relationship between this behaviour and remote covert surveillance detection.

However, the experimenter considered other possible explanations intended to assist future research. Thirty-five percent of participants in the ‘threatening’ area adjusting their clothing within an observation period that lasted an average of just thirty-seconds could be considered a reasonably high number of occasions - and could therefore consequently support the opinion of both the experimenter and the focus group that the ‘threatening’ area did indeed cause more anxiousness than the ‘safe’ area where a lesser percentage of participants adjusted their clothes - but this simply didn’t translate to reporting accuracy. The time of year could also have been an important factor, as whilst conducting the field research in the winter helped eradicate many of the distractions that naturally occur in warmer weather such as ball games, and large outdoor gatherings - this did mean that people were dressed in layers of warm clothing that covered them fully. Perhaps summer attire would have resulted in a higher level of self-consciousness, and may have therefore resulted in more clothing adjustments if individuals felt as though they were under surveillance.

More promising results were found however, when initial analysis of the data showed that 38% of participants in the ‘threatening’ area turned their heads towards the observer, this seemed to the experimenter as though a behavioural indicator of covert surveillance detection may have been found. Especially when turning ones’ head more than 45 degrees for no obvious reason could be considered a very deliberate action in the absence of any apparent stimuli to cause such an intentional movement. Indeed, when a chi-square test of independence was performed to examine the relationship between participants overall correct reports of surveillance and whether they turned their head toward the observer, a significant relationship was found in the predicted direction with participants more likely to turn their head towards the observer when under surveillance.

Interestingly, when participants were not correct in their self-reports, they almost never turned their head (just 12.5% of the time), whilst half of those who did turn their head were correct in their self-reports. The number of times participants turned their heads was not part of the hypothesis and so was not tested, but the experimenter noticed that the individual who turned their head a total of three times did so when they when they were being observed. Again, this could be coincidental - but it could also indicate that some individuals are more sensitive to remote covert surveillance than others; ironically, this is the very idea that the neurological and psychosocial individual differences variables were intended to help predict. Due to the significant and convincing results participants turning towards the direction of observation, the experimenter suggests this measure continue to feature in future research, and that Sheldrake's (2003a) anecdotal evidence for scopaesthesia which suggested that head-turning is a behaviour reported by Special Forces operatives, snipers, security guards and police officers is considered may be replicable under experimental conditions.

The final factor investigated for possible relationships with self-reporting accuracy was environmental. Light levels were taken for each and every field study trial to examine whether this factor shared a relationship with reporting accuracy, as Farnworth et al. (2020) found that brighter light conditions offer individuals safety cues to detect danger in advance, whereas low light conditions may deprive them of such safety cues needed to reduce the sense of threat. A independent samples t-test was therefore conducted to investigate whether light levels had a relationship with participants' overall self-reporting accuracy in the 'threatening' area. The results showed the relationship to be in the opposite of the predicted direction, with the results indicating that participants were more likely to be correct in their self-reports in lighter conditions - this relationship was found to be non-significant however. This result would have alerted the experimenter to possible sensory cues explaining correct self-reports were it not for the pilot studies demonstrating this to be impossible. The outcome was the opposite to what was expected though as the experimenter's prediction that lower light levels would be associated with higher self-reporting accuracy would be logical according to Gordon and King (2002) who state that individuals develop a 'fear of the dark' from childhood. Indeed, if Rachman (1977) is correct that this emerges from environmental factors such as seeing others display fear of the dark, and repeated warnings regarding the dangers of being robbed or kidnapped at night; this fear would remain, or even be exaggerated in adulthood. Such a point is important given that all the participants in the current study were aged 18+. Gordon and King (2002) also argue that a sense of threat from low light conditions is the result of biological preparedness - and as a sense of threat was thought to be an important element of covert surveillance detection following the results of Study One and Study Two, the non-significant outcome related to light levels is surprising. Taken as whole however, the findings related to environmental and behavioural variables partially supported the experimenter's exploratory hypothesis, with 'head turns' and 'walking speed' demonstrating significant results with reporting accuracy - albeit with the latter being in the opposite direction to the that which was predicted.

In summary, whilst of the three measures (self-reports, EDA and task performance) of detection tested in Study One and Study Two, the task performance measure arguably produced the most convincing results, it would appear that this was due to the stress caused by the cognitive task. The literature however would suggest that self-reports would be the most likely to produce positive results following the sheer quantity of studies conducted claiming to have found an effect (Sheldrake, 2003) in relation to scopaesthesia. The field experiment that formed Study Three offered the opportunity to build on the historically reliable self-report measure of surveillance detection, whilst testing the researcher's emerging theory that stress and/or threat is a necessary ingredient for covert surveillance detection to be as efficient as possible.

By having the participants self-report feelings of surveillance in two distinct areas, one of which was deemed to be more threatening than the other - the experimenter could test both the accuracy of their participants' self-reports, along with whether a stressful environment improved their covert surveillance detection abilty. This was indeed the case in two of analyses related to the formal hypotheses. In both instances, it was the threatening area that yielded positive self-report results; the first when self-reporting accuracy was significantly different from chance expectation, and the second when overall self-reporting accuracy was found to be significantly greater in the 'threatening' area compared with the 'safe' area.

Even when the hypothesis was not supported however, as was the case with the experimenter's prediction that covert surveillance self-reporting accuracy will be significantly greater in each area when participants are asked to report surveillance, when compared to non-surveillance - there was still interesting evidence to be found. A significant result would have provided supporting evidence for Colwell et al. (2000), Radin (2004) and Sheldrake (1998, 1999, 2000, 2001, 2005b) who have all argued that whilst they have found participants to be able to detect the presence of a stare, they are less able to detect its absence. Whilst this was not demonstrated to be the case in Study Three, the results were in the predicted direction. More importantly, if the self-reports are investigated for the participants' who were, or were not under surveillance in the 'safe' area - their overall reports could not have been any closer to chance as they were correct literally 50% of the time. However closer inspection shows that they were correct more often when there were under surveillance (see table 26), compared to when they were not. This is not especially convincing due to the small difference (24% vs. 26% respectively), however if the self-reports of the participants who took part in the 'threatening' area are analysed, they were correct overall 69% of time - and actually at a 76% correct surveillance detection rate if surveillance was actually taking place (see table 27). This finding is as Sheldrake would predict, as he found during his research (1998, 1999, 2000, 2001, 2005b) that the majority of participants' correct 'guesses' were made when the individuals were actually being stared at.

Whilst not all every hypothesis related to Study Three found support from the field research conducted, and only one of the predictions was met entirely - taken as a whole, the experimenter felt that the results were valuable in and of themselves, but are important when it comes to informing future research methodology. Between them, the two laboratory studies along with the field study provide an abundance of information which can be used to enhance further research in this field and others, and this will be addressed in the General Discussion. However, an important finding directly related to Study Three that anyone designing similar research should be cautious of is related to the apparent importance of the atmosphere in which the experiment is conducted.

It is perhaps noteworthy that a total of 74 participants walked across the 'safe' area during a relatively short period of time. However, just 26 participants crossed the expanse of the 'threatening' area despite many more hours being spent there by the experimenter and the research assistant. This could be argued to support the researcher's assertion that these two areas were distinctly different surroundings that exuded opposing atmospheres, as it would be expected that individuals would actively avoid an environment they judge to be threatening or dangerous. This would appear to pose an obstacle to studies of this type though, as researching the 'threatening' area took more than quadruple the amount of time and financial resources than that which was dedicated to the research conducted in the 'safe' area - and yet it still resulted in only approximately a third of the data. The experimenter consequently advises that following analysis of the results, such difficult and uninviting atmospheres *should* be sought and used in replications or extensions of the current study - but that the more appropriate the experimental setup, the harder and more expensive the research will be to complete as potential participants actively avoid the area that is ideal for studying the subject of covert surveillance detection.

It should be noted that such difficulties with finding participants within the areas most likely to yield the best results will only be exaggerated if future researchers follow the literature (Gordon & King, 2002) which indicates that despite the non-significant results of Study Three related to light levels, this measure may yet be helpful to improving remote surveillance detection accuracy. The experimenter suggests that perhaps whilst light levels were recorded and used as an environmental variable in Study Three, due to the limitations of the equipment available, the experiment was never conducted in darkness, and that this may account for the non-significant results. This idea is discussed further in the General Discussion. However, future experimenters with the appropriate recording equipment and consideration given to ethical standards could use far lower light levels than was used in the current study to enhance participants' anxiousness as this may create a greater detection accuracy - however, this suggestion comes with the warning that such methodology could make participants even more likely to avoid the intended observation area.

5 GENERAL DISCUSSION

Each of the three studies conducted aimed to find whether remote surveillance detection is possible, which measures and methodology is best suited to capturing this signal should it exist, and which individual differences could be associated with such an ability. The reality of investigating phenomena that may not even exist was always going to be challenging. However, when a thorough review of the literature revealed that even the researchers who claim to have found evidence that scopaesthesia is possible admit that the sense of being stared at is difficult to capture, it was clear that any experimentation focused on something so elusive may yield no results at all.

Whilst over a century of work dedicated to this subject suggests that carefully controlling for sensory leakage may have led to the loss of positive results, the current researcher was determined to eradicate the problems of past studies by ensuring that conventional senses and possible artefacts such as guessing patterns could not account for any apparent evidence of remote surveillance detection. The importance of such considerations had been highlighted by authors such as Marks and Colwell (2000) who have insisted that the opportunity for sensory leakage should be removed, and that all trials should be properly randomised.

Following the considerable attention given to the methodological design and the subsequent pilot studies which demonstrated this to be effective, interesting and potentially useful data was produced that appears to indicate it is possible for individuals to detect covert surveillance via unconventional means. By utilising all three possible methods of signal capture within the same body of research, the experimenter ensured the best possible chance of finding results that may show the ability to sense being watched or listened to exists, and created the opportunity to develop a theory that was driven by comparing and contrasting these differing methods. It was this unique approach that led to the possibility of not only investigating whether scopaesthesia (Sheldrake, 2003) and acoustasthesia (Friday & Luke, 2014) are a real phenomena, but also to examine which situations and circumstances might be necessary to activate or enhance such abilities.

Prior to beginning the experiments, the current researcher hypothesised that participants' correct self-reports would be significantly greater than incorrect self-reports which would indicate that these individuals were consciously aware that they were being watched. However, the first version of the experiment dedicated to self-reports demonstrated that there was no significant difference between participants correct and incorrect self-reports in any surveillance condition. Put another way, the data indicated that participants would have achieved similar results if they were randomly guessing whether or not they were under covert surveillance.

Closer inspection of the self-report data from Study One suggested however that whilst there may not have been a significant difference between participants' correct and incorrect scores, participants' results were in the predicted direction regardless of the surveillance condition (no surveillance, listened to, watched, or watched and listened to). The literature suggests that scopaesthesia is a subtle effect, and even the more convincing evidence claims a success rate of approximately 55% (Sheldrake, 2003) when chance would predict 50%. Whilst the current researcher remained open to the possibility that their research could suggest that critics may be correct to declare that seemingly positive results can be explained by coincidental reactions, confirmation bias, or behaviour driven by normal sensory means - the fact that all of the results were in the predicted direction seemed noteworthy. This encouraged the experimenter to test the self-report measure again, but with an important methodological change as there was a reporting bias evident in Study One, and a power analysis demonstrated that either more participants, or more trials may unveil a previously hidden signal.

Perhaps the self-report related results of Study One should not have been surprising as an accurate staring detection rate of 54.5% was discovered, along with a 54.2% false positive rate which is comparable with the work of Lobach and Bierman (2004) who found that the majority of their participants scored more positively than negatively. Again, this seemed to suggest that there may be a signal to detect, albeit weak and difficult to capture. The problem of reporting bias had also previously been repeatedly identified by Sheldrake (1998, 1999, 2000, 2001, 2005b). Indeed the experimenter's careful analysis of Study One's results showed that participants were significantly more likely to answer positively when asked if they could detect being watched or listened to, and did so 53% of the time. Whilst this bias skewed the results, it was relatively simple to remove as it was likely to have been inadvertently created by the researcher when they introduced the experimental aims to the participants and explained that randomised surveillance would indeed be taking place.

Having learned from the results of the first version of the self-report protocol and identified what may need to be improved upon experimentally, a second study was developed that incorporated methodological changes which addressed these limitations. As finding reliable and attentive participants had already proved difficult and time consuming during Study One, it was decided that statistical power should be improved by doubling the number of trials each participant undertook, rather than increasing the number of individuals who would take part. The researcher's script was also amended to withhold the information that randomised surveillance would take place, so that participants only became aware of this *after* the experiment was finished in an attempt to decrease or eradicate the reporting bias created by priming the participants when they were made aware of this from the beginning of the session. This methodological change had the intended effect, as analysis showed that the reporting bias had been all but eradicated with positive responses overall at a 50.29% response rate which was not significantly higher than mean chance expectation (50%).

Interestingly, all of Study Two's self-report results were again in the predicted direction, and participants' correct self-reports were significantly different to their incorrect self-reports when they were being watched, and when they were watched and listened to. This suggests that participants were not simply guessing whether they were under surveillance when these conditions took place, but may have been sensing a signal. Indeed, participants' total true response rate to mean chance expectation also showed a significant increase, as did their total false response rate. This additional exploratory analysis showed that participants were correct 56.59% of the time in both the 'watched' and the 'watched and listened to' conditions as the work of Colwell et al. (2000), Radin (2004) and Sheldrake (1998, 1999, 2000, 2001, 2003, 2005b) suggested would be the case when they reported a tendency for participants in their studies to know when they were being watched.

There was also a reverse tendency (45.91%) for participants to self-report correctly when no surveillance was taking place, as Sheldrake (2005) would expect based on his claim that the absence of a stare is more difficult to detect than the existence of a stare. When all correct responses were combined, the data showed that participants guessed correctly 6.72% more often than would be expected by chance. This was a highly significant difference which is consistent with Sheldrake's (1998, 1999, 2000, 2001, 2003, 2005b) findings. These results suggest that the methodological amendments implemented between Studies One and Two made an important and discernible difference to the outcome. This was not only important for the current research, but is a finding that could assist future experimenters investigating any topic in which priming could play a role.

Whilst the current researcher admits that positive results were not found under all self-report surveillance conditions, they do suggest that the fact that they were *always* in the predicted direction in *both* laboratory based versions of the self-report protocol is important and unlikely to be coincidental; especially when taken in context with the pattern of the data largely mirroring the findings of others' (Colwell et al,. 2000; Radin, 2004; Sheldrake, 1998, 1999, 2000, 2001, 2003, 2005a, 2005b). Furthermore, even when participants' self-reporting ability was investigated in the real-world setting of Study Three's field experiment to address the possible ecological validity issues associated with the tightly controlled artificial settings of Study One and Study Two, evidence of reporting accuracy remained. This was however limited, and dependent on conditions. As the participants were not aware that the field experiment which formed Study Three was taking place, it was predicted that reporting bias would not be an issue. Analysis of the self-reporting data showed this to be the case, and that that no such bias existed. When looking at how participants' self-reporting accuracy compared with chance expectation (50%), analysis showed that the difference between them was statistically significant in the area deemed to be the most threatening, although this was not the case in the 'safe' area. This partially supported the hypothesis that participants' covert surveillance self-reporting accuracy will be significantly greater than chance expectation within a real-world setting.

The 'threatening' area was again shown to produce significant self-report results when the hypothesis that overall self-reporting accuracy will be significantly greater in the 'threatening' area compared with the 'safe' area was supported. However, covert surveillance self-reporting accuracy was not significantly greater in each area when participants were asked to report surveillance, when compared to non-surveillance as predicted - although closer inspection of the data shows that this is likely to be due to the low number of participants (N = 26). Indeed, reporting accuracy was higher in the surveillance condition (77%) compared to the non-surveillance condition (61.5%) for the threatening condition.

With two of the three formal hypotheses related to Study Three at least partially supported, it was concluded that there may well be a detectable signal which an individual can be consciously aware of. This would appear to be the case in a real-world environment as demonstrated by the field study, but also within a laboratory experiment even when sensory leakage and artefacts are controlled for - so long as reporting bias and statistical power are carefully considered. This supports much of the previous literature; however, this also provided evidence for the experimenters' developing proposal that a degree of stress or danger needs to be present for participants' covert detection ability to manifest. This is suggested due to the positive results of Study Three exclusively relating to the 'threatening' area, rather than the 'safe' area, and is further supported by the task performance results of Study One and Study Two.

The current researcher analysed the task performance data of the laboratory based research relating to Study One and Study Two and discovered unexpected, but compelling evidence that this method of remote surveillance detection appears to be the most reliable having produced similar results in both experiments. Initially this method was included based on the idea that people tend to behave and perform differently when alone compared to when they are in the presence of others (Griffin & Kent, 1998), which led to the idea that if people do so under conventional observation, the same may be true when the observation is covert. However, due to the lack of evidence in the literature to support it - the hypothesis which predicted that the time taken for participants' to accurately complete the Stroop task will be different under covert surveillance, compared to when not under covert surveillance was based somewhat on a logical leap.

When the self-report results of Study One indicated no effect, or a weak signal at best (as the positive results following the methodological change applied to Study Two had not yet been discovered at this early stage of the research), the experimenter theorised that task performance results were unlikely. The evidence up until this point had suggested individuals were not consciously aware of surveillance, and the EDA measure had thus far indicated that they were not unconsciously aware either. It therefore followed that task performance should not be a sufficiently sensitive measure of covert surveillance detection as the participants seemed unable to 'feel' being watched or listened to under the tightly controlled laboratory conditions either consciously or physiologically. In fact, the experimenter continued with the task performance protocol related to Study One more for thoroughness and integrity having stated their hypotheses, rather than through a realistic expectation of positive results. However, despite their own scepticism, the experimenter found task performance not only to be a measurable indicator of covert surveillance detection - but that it produced the only significant covert surveillance detection results of Study One.

Paired t-tests showed that when comparing the means of participants' task performance reaction times, two out of the three possible combinations were significantly different from one another. Specifically, participants took longer to complete the task in the 'watched' and the 'watched and listened to' conditions when compared to the 'no-surveillance' condition. These results appeared to suggest that covert surveillance can be detected via non-conventional means, and that this detection can delay an individual's decision-making ability as would be expected according to the Yerkes-Dodson law (Yerkes & Dodson, 1908). This continued to be true in Study Two's version of the task performance experiment, however this was only the case for one surveillance condition combination rather than two. This time, only the 'watched and listened to' condition had a statistically significant effect on participants' reaction times when compared to the 'no surveillance' condition.

The researcher admits that the second partial replication of the task performance related element of the experiment produced less convincing results than the initial version, but suggests it is important to note which of the surveillance condition means were significantly different from one another, as this was the combination of 'no surveillance' compared with 'watched and listened to'. These two conditions should theoretically demonstrate the greatest difference in means as according to the literature (and indeed, the current studies' results), the 'no surveillance' condition should be the least detectable as there is no signal to detect. Conversely, the participants being watched and listened to should have experienced the most interference to their decision making due to the effects of scopaesthesia, compounded by any additional signal caused by being listened to. Study Two's task performance protocol may have produced fewer positive results than Study One, but the researcher suggests that this is meaningful - especially in relation to the idea that adding the 'listened to' element to scopaesthesia increases its detectability, as if this was not the case - there is no reason why the 'watched and listened to' condition should cause more interference to reaction times than the 'watched' condition. This would appear to suggest that acoustathesia may not exist in its own right independently of scopaesthesia, but that a combination of these two signals results in the most detectable surveillance condition.

As well as providing limited evidence that being listened to is to some extent remotely detectable, the experimenter also suggests that comparisons between the task performance results of Study One and Study Two provides yet more support for the proposal that stress and anxiety increase the effects of remote surveillance detection. This cautiously stated conclusion is based on the less convincing results of Study Two when compared to Study One, as the only methodological change related to this protocol was that participants were not told that covert surveillance would take place on the second occasion. This single amendment resulting in differing results could be argued to demonstrate that the participants' knowledge that they would be watched and listened to may have caused the stress that could have led to the positive results extending to two of the three surveillance conditions in Study One, rather than just one surveillance condition in Study Two.

As a possible means of covert surveillance detection, physiological response measured by changes in participants' EDA is in some ways the opposite of self-reports; the latter requires conscious awareness of the surveillance and deliberate action, whereas the former could be measured without the individual even knowing their body had reacted to the differing surveillance conditions. With Lobach and Bierman's (2004) work suggesting that EDA would provide measurable results, it was hypothesised that the current researcher would find significant differences between participants' EDA compared with their baseline depending on whether they were under no surveillance, being listened to, being watched, or being watched and listened to. This was however, not the case.

Study One showed participants overall EDA to be 1.2% lower in the 'watched and listened to' condition than the 'no surveillance' condition. This difference was not statistically significant - yet it was the most convincing of the surveillance condition comparisons. Whilst the obvious conclusion to draw would be that psychophysiology may not be a measurable detection method for indicating whether or not an individual is unconsciously aware of covert surveillance, it was considered that the experimenter's welcoming introduction to the participant who would be providing the surveillance along with their introduction to their counterpart may have had a calming effect for some, and an arousing effect for others. This idea was based on the work of Braud et al., (1993a, 1993b) who showed that remote observation can create opposing reactions - if this were the case, the response of some participants could have neutralised the responses of others.

It was thought that perhaps the methodological change applied to Study Two's EDA protocol may affect the results, as by not introducing the observed participant to the individual who would be observing them - the experimenter considered that this may address possible issues caused by such an introduction whereby the meeting could cause either a calming or an arousing effect in participants. Omitting the introduction was a methodological by-product of not telling the participants that randomised surveillance would occur - an amendment originally introduced in an effort to avoid self-reporting bias. As physiological response is not conscious or voluntary - participants' reporting bias was not applicable to the EDA protocol, however, the experimenter surmised that the lack of expectation regarding being under surveillance may play an important role in this second version of the EDA element of the experiment. This however was not demonstrated to be the case, and there were again no significant differences between participants EDA in the surveillance conditions and the control (no surveillance) condition.

The failure to capture any meaningful changes in participants' EDA in either version of the experiment was disappointing. As the current researcher had expected that this measurement would not be susceptible to the reporting bias associated with self-reports as demonstrated in Study One, they had therefore considered that EDA may provide a way to measure surveillance detection without the participant needing to make conscious decisions and so may yield positive results. This method of measurement did come with a warning though via the literature as Schmidt and Walach (2000) stated that it can be prone to artefacts, and Lobach and Bierman (2004) reported a seemingly genuine signal difference of only 1%. Indeed, a standard deviation of 13 - 14% was found during Study One indicating that this may well be the case, and that the weak signal could be drowned out by the noise inherent in this noisy process.

Whilst EDA as an independent measure of remote surveillance detection yielded only non-significant results, this measure when combined with another provided a very interesting outcome. Whilst considered unlikely to make a difference at the time, further thought related to EDA as a measure of surveillance detection not only proved worthwhile - but could actually be considered one of the major findings of the entire series of experiments. The current researcher considered the convincing outcome of the task performance experiments in Study One and Study Two, and theorised that a feasible explanation for this may be that the measure which resulted in the most replicable results could be due to the participants 'need' to detect the surveillance as they were being tested. As an increase in EDA is known to be a direct response to stress (Silverthorn et al., 2009), it was proposed that perhaps with no need to detect whether they were under surveillance in the EDA protocols during which all that was required of the participants was to sit still and read bland text aloud - they never detected being watched and/or listened to, as it was simply unnecessary to do so.

The idea that ‘threat’ or ‘necessity’ is required for surveillance detection to manifest is logical if Sheldrake’s (2003a) anecdotal evidence for scopaesthesia offered by Special Forces operatives, snipers, security guards and police officers is considered. Sheldrake claims that these professionals report that they are able to sense when they are being watched, and in every career example listed, they would *need* to know when this is taking place. It could therefore be extrapolated that perhaps without the necessary stimuli, individuals’ extrasensory ability may be dormant - but is activated under certain specific circumstances. Perhaps in the absence of a situation that matters to the person such as danger, embarrassment, personal gain, or in this case being tested - the individuals will not ‘react’, as they do not need to.

Based on this line of enquiry, the researcher conducted the EDA experiment again. This time though, they asked all participants to complete the Stroop Test (Ridley, 1935) used in the task performance experiments *whilst* their EDA was recorded. This hybrid experiment resulted in two of the three possible surveillance condition comparisons with baseline (no surveillance) being statistically significant. Participants’ EDA increased when they were being watched, and the when they were being watched and listened to. The current researcher tentively suggests that this supports their developing theory that the ability to detect surveillance could be a real phenomenon, but that the necessity to detect is required for the ability to be accessed. Put simply, it needs to ‘matter’ to the individual whether they are under surveillance or not for them to become aware of it.

When it came to the individual differences under investigation, it was hypothesised that these variables would share a relationship with covert surveillance detection should it exist as there is a wealth of evidence in the literature to support that such patterns would be likely - whether this be with belief and experience (Irwin & Green, 1999; Kennedy & Kanthamani, 1995; Kohr, 1980; Lange & Houran, 1997; Lawrence, 1993; Lindeman & Aarino, 2006; McClenon, 1994; Palmer, 1979; Ring, 1984; Zdrenka & Wilson, 2017), temporal lobe lability (Britton & Bootzin, 2004; Irwin, 2009; Luke et al., 2013; Wiseman & Greening, 2005), or Schizotypy (Barnes & Gibson, 2013; Friday and Luke, 2014; Irwin & Green, 1999; Mason et al., 2005; Mathijsen, 2016; McCreery & Claridge, 1995; Simmonds-Moore, 2010). The literature review revealed though that these individual differences related to forms of anomalous experiences that did not extend to remote surveillance detection, and so the investigation of these factors was exploratory.

Whilst generally, most variables were shown to not be related to detection accuracy more often than they were; there were some interesting findings initially. Examples include Pearson's correlation analysis demonstrating that participants' CPES scores shared a relationship with task performance times in the 'watched and listened to' condition, and multiple regression analysis showing that Introvertive Anhedonia and CPES scores both contributed significantly to the model whilst participants were being watched and listened to. However, even in instances where the same individual difference was demonstrated to have a relationship with various detection measures, there was little consistency across the various protocols and studies - this may possibly be explained, at least in part, by multiple analyses. For this reason, Bonferroni corrections were applied and any relationships that were not still significant following this analysis were discounted and not considered meaningful. It is therefore recommended that researchers should carefully consider the inclusion of these individual differences in future research regarding scopaesthesia, or find possible new ways of measuring these factors.

Whilst the experimenter always intended including the well-researched factors of anomalous experience and belief in every version of the experiments due to the seemingly reliable association between them and phenomena which cannot be explained via conventional science, it had been considered whether the questions included in the experience and belief subscales of Gallagher et al.'s AEI Inventory (1994) alone would be relevant enough. The experimenter therefore decided to use these predefined questions, as well as including two additional questions; 'do you believe it is possible to detect being watched' and 'do you believe it is possible to detect being listened to' which were added to the belief subscale to investigate whether specificity was an important consideration. An equivalent two questions were also added to the experience subscale. Analysis of the AEI scores revealed there were no significant relationships found between AEI Experience in Study One or Study Two, regardless of the detection method used. Whilst AEI Belief initially appeared to share a significant relationship with the 'watched and listened to' surveillance condition in Study One specifically if the additional remote detection questions were included, this relationship did not remain after a Bonferroni correction was applied to account for multiple analyses. It was therefore concluded that the anomalous experience and belief variables shared no meaningful relationship with surveillance detection accuracy, regardless of whether the additional specific questions were included or not.

Whilst AEI Experience and AEI Belief were demonstrated to show no significant relationships with covert surveillance detection, Schizotypy was measured by Mason et al.'s (2005) O-LIFE inventory. This was broken down into the four sub-scales of Unusual Experiences, Cognitive Disorganisation, Introvertive Anhedonia, and Impulsive Nonconformity. The Unusual Experiences and Introvertive Anhedonia subscales initially showed promise of a correlation with self-reporting in Study One, but these did not survive Bonferroni correction and neither returned significant models when explored via regression analysis. Similar was true of the Task Performance protocol when investigated via Pearson's correlation analysis. Yet the multiple regressions revealed Introvertive Anhedonia contributing significantly to the model in relation to the task performance protocol when participants were being watched and listened to, and in relation to Impulsive Nonconformity when the participants were under no surveillance. Neither of these subscales remained significant in Study Two though in any of the experimental protocols.

Temporal lobe lability was included as a predictor variable, and was measured by Persinger and Makarec's (1993) CPES inventory. Significant correlations related to self-report scores in Study One did not survive Bonferroni correction, however this factor was found to have a statistically significant correlation with task performance when participants were watched and listened to. The CPES factor was also shown to contribute significantly to the model following multiple regression analysis. This potentially exciting finding was not replicated in Study Two though.

Overall, the search for a relationship between neurological and psychosocial individual differences and covert surveillance detection was disappointing, with the task performance protocol shown to be the only measure to produce more than one significant relationship. A significant correlation was found between task performance and participants' CPES scores whilst being watched and listened to, and this same factor contributed significantly to the model, along with Introvertive Anhedonia under the same surveillance condition following multiple regression analysis. When Impulsive Nonconformity was also found to contribute significantly to the model when participants were under no surveillance - this measure of remote detection appeared to provide the clearest insight into which of the neurological and psychosocial individuals under investigation may be associated with covert surveillance detection ability. These relationships were therefore expected to be demonstrated again via the task performance protocol in Study Two, as whilst there was a methodological amendment - this was related to the self-report protocol.

The methodological change was not designed to affect the task performance protocol, and so the task performance protocol in Study Two was initially thought of as a replication and extension to see if the significant differences between participants' reaction times would be duplicated. To some extent, a similar effect on the time it took participants to complete the Stroop Task was found, and so the single methodological amendment of not informing participants that covert surveillance would be taking place did not eradicate the effectiveness of task performance as a measure of detection, but the relationships between the individual differences variables disappeared. There was unfortunately no consistency across either the detection measurement type (i.e., self-report, EDA, task performance), or the conditions (no surveillance, listened to, watched, and watched and listened to) rendering the individual differences variables less relevant than predicted.

The relevance of these neurological and psychosocial variables was extrapolated from previous psi research, the majority of which involved less passive experiments that demanded more effort and intent from the participants than the current covert surveillance based studies. This may well be an important difference, and could perhaps explain the variation in outcomes between the current research and other studies involving psi. Whilst these neurological and psychosocial factors revealed little association with covert surveillance detection in the laboratory experiments - the individual differences exploratory hypothesis related to Study Three was even more disappointing when despite a focus group indicating it would not be the case - so few of the field study participants provided their individual differences measures that they could not be meaningfully analysed.

The environmental and behavioural variables were not reliant on the participants completing the study though. Light levels were captured by the research assistant at the time the trial took place, and the behavioural changes assumed to be possible indicators of covert surveillance detection such as participants' walking speed, and whether they turned their heads towards the observer, adjusted their clothing, or paused whilst walking were recorded for later analysis. Only one of these factors provided significant results in the predicted direction. In fact, the independent judges agreed that not a single participant paused at all within the designated observation areas, and analysis showed that participants adjusting their clothing had no significant relationship with reporting accuracy. However, participants correctly reporting surveillance tended to walk slower than those who incorrectly reported surveillance, which was opposite to the experimenter's prediction. In the absence of directly related literature, the experimenter based the prediction that participants correctly reporting surveillance would quicken their pace on previously established research regarding individuals' motivation to avoid uncomfortable situations such as darkness (Gordon & King, 2002), but the reverse appeared to be the case. Head turns were found to have a significant relationship with correct self-reporting in the predicted direction though, with half the participants who turned towards the observer being correct in their surveillance reports. This finding is particularly interesting as this is the most often used behavioural response measure, and supports Sheldrake's (2003a) anecdotal evidence from Special Forces operatives, snipers, security guards and police officers who report repeated instances of their target turning to face them in response to being stared at.

Possible explanations for why certain behavioural measures of surveillance detection may have not been significantly related to correct self-reports include their lack of subtlety, the absence of real threat, and the season in which the field experiment took place. For example, walking speed is on a continuum, with the participants taking anywhere between 23 to 40 seconds to pass through the 'threatening' area allowing analysis of changes - whereas pauses were dichotomous, with the judges being asked to record a pause if the participant came to a complete standstill. Perhaps such a behavioural change was too dramatic, especially in a situation in which participants were by definition familiar with the area as they had to be either university students or staff. Ironically, if this were the case, it would mean it was the ethical limitations that made this measure invalid. This line of thinking would indicate that future researchers should design ethically acceptable ways of increasing participants' sense of threat or danger.

The measurement of light levels was introduced to test the idea of threat affecting detection ability further. This idea was supported by the literature, as Farnworth et al. (2020) identified that bright light acts as an indirect visual cue to potential predation and that bright light conditions offer safety cues to detect danger in advance - conversely, low light conditions may deprive individuals of the safety cues which would reduce their sense of threat. This 'fear of the dark' is thought to stem from childhood (Gordon & King, 2002) and so if 'threat' may be an essential element for covert surveillance detection to exist - light levels were thought to be an ethical substitute for genuine threat and anxiety that is reasonably easy for future experimenters to control in a laboratory based environment, and simple to measure in a field experiment.

Despite the non-significant results, the light level variable could, with different methodology, still produce positive results in future studies as the current field experiment had limitations regarding this variable due to the areas and equipment used. Whilst light levels were recorded in conditions which ranged from full sunlight to dusk - participants could not be clearly seen on the recordings if the experiment continued after dark. Future experimenters could take this into account when designing the study and factor in this obstacle. Indeed, anyone wanting to replicate the study could do so without changing the methodology by simply using a night vision camera. Even this would not be necessary if future researchers were not measuring changes in participants' behaviour, as the use of the recordings were largely included for later analysis by independent judges - something that would not be necessary for an experimenter interested only in self-reporting accuracy rather than the behavioural changes, especially as they brought limited results.

The practicalities of future research should be considered, as completion of all three studies took the current researcher far longer than anticipated due to methodological difficulties encountered with such technically challenging experiments. However, through an open-minded and scientific approach, the researcher followed where the evidence led and consistently adapted their methodology, which was always dictated by the findings; in doing so, the researcher answered many of the questions the experiments were created to address, and through exploratory analysis - some that had not even been asked.

Overall, the researcher found their results to be both expected *and* surprising having produced data that broadly reflects the existing literature such as the self-report detection rates, whilst some of the research opposes that which preceded it, such as the EDA experiments. Evidence was also found to indicate that under-researched measures of surveillance detection such as task performance may be an apparently replicable method of investigation - this discovery was not only important in its own right, but also formed the basis of the experimenters' proposal regarding necessity, threat or stressful situations promoting improved covert surveillance ability. This has interesting implications for future research that may enhance researchers' methodology; for example, embarrassing situations may cause participants to want to know whether they are under surveillance, as could scenarios in which they have the opportunity to cheat or avoid a monotonous task by discontinuing it when they believe nobody is watching or listening to them.

The literature supports this idea as stress and anxiety can impact performance by negatively affecting processing efficiency and reducing attentional resources (Broadbent et al., 2019). An increase in stress and anxiety can create a 'cognitive overload' of limited cognitive resources that are competing with other processes (Sariggianadis et al., 2020), which can impact attention, focus and performance. Anxiety not only influences cognitive overload, but it can also influence physiological changes such as elevated heart rate and increased cortisol levels, whereby high cortisol levels generated by anxiety/stress are associated with decreased performance, but low to moderate levels of elevated cortisol may improve performance (Schmidt et al., 2020). Therefore, participants need to achieve an optimal level of arousal/anxiety to generate physiological reactions that are associated with improved performance (Weinberg, 2011). In addition, some researchers argue that achieving an optimal level of attentional control can also help participants achieve an optimal level of anxiety that aids performance (Ducrocq et al., 2017), as a loss of attentional control due to high stress/anxiety may deprecate performance. Increased stress and anxiety levels can also impact subjective experiences and perceptions of one's environment, whereby anxiety can lead participants to underestimate the duration of temporal intervals (Sariggianadis et al., 2020). Therefore, anxiety may make time appear to go faster which indicates that increased anxiety/stress may impact perceptions of one's environment and influence one's subjective experiences which may impact performance.

It is the potential discovery concerning the possible importance of stress and anxiety in relation to scopaesthesia (Sheldrake, 2003) that the author suggests may be one of the most important contributions their body of work may offer future researchers in safe and ethically responsible ways. Previous studies have suggested that anxiety can be successfully induced through virtual reality (Mostajeran et al., 2020), whereby creating the pressure of a large audience through virtual reality has been shown to increase social anxiety as measured by changes in heart rate. The use of virtual reality and augmented reality have been found to increase anxiety in experimentally controlled environments, and can be utilised to trigger gradual rises in anxiety (Tsai et al., 2018; Yin et al., 2020). However, Tsai et al. (2018) point out that a virtual reality environment generates a greater increase in anxiety than an augmented reality environment. The use of virtual reality to induce threat or anxiety also presents neurological implications, whereby researchers can monitor the activation of specific brain regions in response to environmental stimuli generated by virtual reality (Goodman et al., 2020).

Whilst setting out to investigate whether it is possible for individuals to detect covert surveillance via unconventional means - through a willingness to consider ideas beyond the initial hypotheses, the author cautiously suggests that not only has evidence of covert surveillance detection been found, but also the specific situation in which it is most likely to occur may have been discovered. The theory that the presence of threat or necessity makes it more likely that an individual can access their ability to detect covert surveillance is not only important in and of itself, but should also be considered by researchers in future studies.

The author acknowledges that real and actual danger cannot and should not be introduced within experiments intended to further investigate this idea, however it is suggested that situations in which the participants feel judged (as evidenced by the experiments related to task performance), embarrassed, or self-conscious may prove vital in this research area via the use of virtual reality. The author would also encourage the inclusion of the sense of being listened to, as whilst there was indeed some limited evidence within the results to support the idea that acoustasthesia (Friday & Luke, 2014) may exist, the data seems to suggest that listening to an individual while watching them results in a condition more detectable than if scopaesthesia (Sheldrake, 2003) is measured independently.

The author would like to highlight that with replication and robust supportive research, these discoveries are potentially useful outside of the laboratory setting and could be applied to real-world situations in which surveillance plays a role. These could include everyday examples such as occupational and education settings in which employees and students are monitored for performance, as being watched has been found to increase distress intensity (O'Sullivan et al., 2020) and makes the individual being observed feel like they have lost control over their privacy (Esmark et al., 2017). The sensation of being watched increases during high arousal situations such as examinations (Hesslinger et al., 2017) which suggests that the sensation of being watched during exams may elevate feelings of distress and, as a result, be detrimental to exam and academic performance.

If the sense of being watched is indeed damaging to academic performance, taking away such surveillance should be seriously considered. A seemingly obvious defence of watching students taking exams is to reduce cheating, however previous research has found that the sensation of being watched during exam situations does not significantly affect cheating (Pfattheicher et al., 2019), which is supported by research that suggests being watched does not increase prosocial behaviour or self-referential processing (Canigueral & Hamilton, 2020). Conty et al. (2010) argue that being watched can present consequences such as increased distraction and loss of attention, which highlights that being observed during high pressure situations such as examinations may be highly detrimental to performance, while offering no benefits to the prevention of cheating behaviour.

There are also more extreme clinical applications to consider such as when a symptom of a disorder is paranoia, which may turn out to be based on the sensing of attention directed at the patient, or the more obvious but important use of covert surveillance detection information by agents involved in matters of security. The latter could include everything from international terrorism, to personal security. Indeed, the feeling of being watched has been found to establish real-world implications for cyber-security and reduced information security violations, whereby the perceived presence of an onlooker can generate negative emotions (such as guilt or shame), in potential cyber-security violators and so reduces the likelihood of information security violations, due to perceived consequences (Farshadkhah et al., 2021). Pfattheicher and Keller (2015) support the idea that the perceived presence of others increases prosocial behaviour (Nakagawa et al., 2019) and elevates public self-awareness. However, other real-world implications for scopaesthesia (Sheldrake, 2003) extend into an individual's sense of safety while using social media, online privacy and privacy rights (Bogue, 2019). It is sincerely hoped that this body of research will be shown to be ecologically valid, and will contribute in a meaningful way to these areas.

The results and ideas generated from the current research may also warrant revaluation of past remote staring experiments, as the importance of the reporting bias unveiled in Study One and Study Two may have implications for previous positive self-reports (Colwell et al., 2000; Peterson, 1978; Sheldrake, 2000). Similarly, it would be interesting to know whether Colwell et al. (2000), Peterson (1978), Sheldrake (2000), or Williams (1983) had deliberately or accidentally introduced an element of threat or stress when they gained positive results as measured by EDA in their remote observation experiments. It would also be fascinating to learn more about to what extent the behavioural differences research (Chen, 1937; Cottrell et al., 1968; Dashiell, 1935; Platt et al., 1967; Travis, 1925; Triplett, 1898) which initially instigated the experimenter's theory that an individuals' actions may be affected by covert surveillance were stressful or threatening. The value of the data would also extend scopaesthesia studies, and has implications for all variations of ESP (Rhine, 1934) including telepathy, intuition, psychometry, clairvoyance, clairaudience, and their associated trans-temporal operations such as retrocognition or precognition.

Previously psi research has been largely divided into proof-oriented experiments designed to evaluate whether there is a psi effect, and process-oriented research conducted by those who assume the existence of psi and instead endeavour to understand how and why it manifests (Cardeña et al, 2017). The current researcher attempted to assist with the latter through measuring which psychosocial and neurocognitive factors are characteristic of individuals who perform well in laboratory based psi research, but to some extent failed to do so when there were no consistent results relating to the individual differences factors under investigation across the experiments. This finding highlighted that measures cannot simply be extrapolated from apparently similar research, and by association perhaps provides evidence that the various strands of psi cannot, and should not be treated as the same phenomena simply because they feature an element of ESP. Future researchers may therefore find that alternative psychosocial and neurocognitive factors could be associated with surveillance detection, as perhaps scopaesthesia is not related to the same individual differences variables as other ESP related abilities.

The current researcher also suggests that elements of their research could be considered by others wanting to investigate anomalous phenomena, as whilst the aforementioned lesson regarding assumed extrapolation of variables was accidental, many of the deliberate approaches proved beneficial and could be applied in future research on related topics. For instance, by being honest and realistic regarding the ecological validity of their laboratory based work and investigating whether the apparent phenomenon could be replicated in a real-world setting, the researcher produced results far more convincing than if a multidisciplinary approach had not been adopted. Similarly, by not relying on just one method to capture the signal should it manifest, the researcher not only gave themselves the best possible chance of doing so - but also allowed themselves the opportunity to cross reference the data. This resulted in the ability to contrast and compare detection measures in a way that would have otherwise been impossible.

The willingness of the researcher to consider serendipitous or seemingly redundant data rather than discard it could also be adopted by fellow investigators, as it was this approach that led to one of the most valuable findings when the EDA results that seemed to be irrelevant became an essential part of discovering the apparent value of stress and anxiety. Had these initially disappointing physiological response readings been ignored instead of being used as a basis for evolving the methodology used in Study Two, the important finding that stress is a necessary component for covert surveillance detection to manifest could have been missed altogether. In fact, the evolution of the experimental design was perhaps the main reason meaningful discoveries were made, and this open-mindedness should be encouraged in future experiments, just as it has been since the topic was first investigated by the founders of psychology, who strived for a comprehensive and transparent discipline. The author humbly hopes that their contribution to this under-researched field is of value in its own right, but that it also becomes a stepping-stone for future researchers who can build upon the results and emerging theories to further a joint quest towards uncovering the potential of the human mind and its capabilities.

6 REFERENCES

Alcock, J. (2003). Give the null hypothesis a chance: reasons to remain doubtful about the existence of psi. *Journal of Consciousness Studies, 10,* 29-50.

Alcock, J. (2010). Attributions about impossible things. In S. Krippner & H. L. Friedman (Eds.), *Debating psychic experience: Human potential or human illusion?* Santa Barbara, CA: Praeger.

Alvarez, L. W. (1965). A pseudo experience in parapsychology. *Science, 148,* 1541-1965.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5^{th} ed.).

Anderson, M. C. (2012). Non-intentional behavioural responses to psi: Hidden targets and hidden observers, Unpublished MA thesis, University of Edinburgh.

Andrews, S. (2004). The human connection institute. Retrieved 2005 from http://www.connectioninstitute.org/index.htm.

Atkinson, A. P. (2005). Staring at the back of someone's head is no signal, and a sense of being stared at is no sense. *Journal of Consciousness Studies, 12*(6), 50-116.

Bach-y-Rita, P. (1972). *Brain mechanisms in sensory substitution*. Academic Press, New York.

Bach-y-Rita, P. (1995). *Nonsynaptic diffusion neurotransmission and late brain reorganization.* Demos-Vermande, New York.

Bach-y-Rita, P., Tyler, M. E., & Kaczmarek, K. A. (2003). Seeing with the brain. *International Journal of Human-Computer Interaction, 15*(2), 285-295.

Bak, M., Myin-Germeys, I., Delespaul, P., Vollebergh, W., de Graaf, R., & van Os, J. (2003). Do different psychotic experiences differentially predict need for care in the general population? *Comprehensive Psychiatry, 46*, 192-199.

Baker, I. (2005). Nomenclature and methodology. In A. Freeman (Ed.), *Sheldrake and his critics: The sense of being glared at*. Exeter, UK: Imprint Academic.

Banks, I. (2017). The NHS Direct healthcare guide. Retrieved from www.healthcareguide.nhsdirect.nhs.uk

Barnes, K., & Gibson, N. (2013). Supernatural Agency: Individual Difference Predictors and Situational Correlates. *International Journal for the Psychology of Religion*, *23*(1), 42-62.

Bell, V., Halligan, P., & Ellis, H. (2007). The psychosis continuum and the Cardiff anomalous perceptions scale (CAPS): Are there multiple factors underlying anomalous experience? *European Psychiatry*, *22*(1), Published abstract.

Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin*, *115*, 4-18.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology, 100,* 407-425.

Bem, D. J., Palmer, J., & Broughton, R. S. (2001). Updating the Ganzfeld database: A victim of its own success? *Journal of Parapsychology, 65,* 207-218.

Bem, D., Tressoldi, P., Rabeyron, T., & Duggan, M. (2015). Feeling the future: A meta-analysis of 90 experiments on the anomalous anticipation of random future events (version 2). *Research, 4,* 1188.

Bennett, P. J., & Cortese, F. (1996). Masking of spatial frequency in visual memory depends on distal, not retinal, frequency. *Vision Research, 36*, 233-238.

Bennion, M. (2018). Eyes in the back of your head? Belief in scopaesthesia and its relation to paranormal belief and cognitive-perceptual personality measures. Unpublished thesis, Manchester Metropolitan University.

Bentall, R., & Slade, P. (1985). Reality testing and auditory hallucinations: A signal-detection analysis. *British Journal of Clinical Psychology*, *24*, 159-169.

Bentall, R. (1990). The illusion of reality: A review and integration of psychological research on hallucinations. *Psychological Bulletin*, *107*, 82-95.

Bentall, R. (2000). Research into psychotic symptoms: are there implications for parapsychologists? *European Journal of Parapsychology*, *15*, 79-88.

Burbeck, C. A. (1987). Locus of spatial-frequency discrimination. *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, *4*, 1807-1813.

Berenbaum, H., Kerns, J., & Raghavan, C. (2000). Anomalous experiences, peculiarity, and psychopathology. In E. Cardeña, S. J. Lynn & S. Krippner (Eds.), *Varieties of anomalous experience: Examining the scientific evidence*, pp. 25-46. Washington DC: American Psychological Association.

Bering, J. M. (2006). The folk psychology of souls. *Behavioural and Brain Sciences, 29*, 453-46.

Blackmore, S. J. (1984). A postal survey of OBEs and other experiences. *Journal of the Society for Psychical Research, 52*, 225-244.

Boden, M., & Berenbaum, H. (2004). The potentially adaptive features of peculiar beliefs. *Personality and Individual Differences, 37*, 707-719.

Bogue, R. (2019). Seeing and Being Seen. *Hedgehog Review, 21*(1), 86-95.

Bond, C. F., & Titus, L. J. (1983). Social Facilitation: A meta-analysis of 241 Studies. *Psychological Bulletin*, *94*, 265-292.

Bösch, H., Steinkamp, F., & Boller, E. (2006). Examining psychokinesis: The interaction of human intention with random number generators - A meta-analysis. *Psychological Bulletin, 132,* 497-523.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Society for psychophysiological research ad hoc committee on electrodermal measures. *Psychophysiology, 49,* 1017-1034.

Braud, W. G. (2003). Distant mental influence, its contributions to science, healing and human interactions. Charlottesville: Hampton Roads Publishing Company.

Braud, W., & Shafer, D. (1993). Further studies of autonomic detection of remote staring: Replication, new control procedures, and personality correlates. *The Journal of Parapsychology, 57,* 391-409.

Braud, W. G., Shafer, D., & Andrews, S. (1990). Electrodermal Correlates of Remote Attention: Autonomic Reactions to an Unseen Gaze. *The Parapsychological Association 33rd Annual Convention Proceedings of Presented Papers, 33,* 14-28.

Braud, W., Shafer, D., & Andrews, S. (1993a). Reactions to an unseen gaze (remote attention): A review, with new data on autonomic staring detection. *Journal of Parapsychology, 57,* 372-390.

Braud, W., Shafer, D., & Andrews, S. (1993b). Further studies of autonomic detection of remote staring: Replications, new control procedures, and personality correlates. *Journal of Parapsychology, 57,* 391-409.

Bressan, P., Kramer, P., & Germani, M. (2008). Visual attention capture predicts religiosity. *Cortex, 44,* 1299-1306.

Britton, W. B., & Bootzin, R. R. (2004). Near-death experiences and the temporal lobe. *Psychological Science: A Journal of the American Psychological Society, 15*(4), 254-258.

Broadbent, D., Gredin, N., Rye, J., Williams, A., & Bishop, D. (2019). The impact of contextual priors and anxiety on performance effectiveness and processing efficiency in anticipation. *Cognition and Emotion*, *33*(3), 589-596.

Brod, J. (1997). Creativity and schizotypy. In G. Claridge (Eds.), *Schizotypy: Implications for illness and health*, (pp. 274-298). New York, NY: Oxford University Press.

Broughton, R. S. (1987). Publication policy and the Journal of Parapsychology. *Journal of Parapsychology*, *51*, 21-32.

Broughton, R. S. (1988). If you want to know how it works, first find out what it's for. In D.H. Weiner & R.L. Morris, (Eds.), *Research in Parapsychology, 1987* (pp.187-202). Metuchen, NJ: Scarecrow Press.

Broughton, R. S. (1991). *Parapsychology: The controversial science*. London Sydney Aukland Johannesburg: Rider.

Broughton, R. S. (2015). Psi and biology: An evolutionary perspective. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 139-148). Jefferson, NC: McFarland.

Broughton, R. S., and Perlstrom, J. R. (1992). PK in a competitive computer game: A replication. *Journal of Parapsychology, 56*, 292-305.

Burns, J. E. (2005). Detection of staring - psi or statistical artefact? *Journal of Consciousness Studies*, *12*(6), 71-75.

Brugger, P., & Mohr, C. (2008). The paranormal mind: How the study of anomalous experiences and beliefs may inform cognitive neuroscience. *Cortex, 44,* 1291-1298.

Burch, G., Hemsley, D., & Joseph, M. (2004). Trials to criterion latent inhibition in humans as a function of stimulus pre-exposure and positive schizotypy. *British Journal of Psychology*, *95*, 179-196.

Burch, G. S., Steel, C., & Hemsley, D. R. (1998). Oxford-Liverpool inventory of feelings and experiences: Reliability in an experimental population. *British Journal of Clinical Psychology, 37,* 107-108.

Cañigueral, R., & Hamilton, A. (2019). Effects of being watched on self-referential processing, self-awareness and prosocial behaviour. *Consciousness and Cognition*, *76*, 102-104.

Cardeña, E. (2010). A look at the potential future of parapsychology with help from a regression to the hypnotic past. *Journal of Parapsychology 74,* 15-30.

Cardeña, E. (2014). A call for an open, informed study of all aspects of consciousness. *Frontiers in Human Neuroscience, 8,* 17.

Cardeña, E. (2015). The unbearable fear of psi: On scientific censorship in the 21st century. *Journal of Scientific Exploration*, *29*, 601-620.

Cardeña, E. (2018). The Experimental Evidence for Parapsychological Phenomena: A Review. *American Psychologist*. Online publication. http://dx.doi.org/10.1037/amp0000236.

Cardeña, E., Lynn, S. J., & Krippner, S. (2017). The psychology of anomalous experience: A rediscovery. *Psychology of Consciousness, 4,* 4-22.

Cardeña, E., Palmer, J., & Marcusson-Clavertz, D. (2015). *Parapsychology: A handbook for the 21st century*. Jefferson, NC: McFarland.

Carpenter, R. H. S. (2005). Does scopesthesia imply extramission? *Journal of Consciousness Studies*, *12*, 76-77.

Carpenter, J. (2012). *First sight: ESP and parapsychology in everyday life*. Lanham, MD: Rowman & Littlefield.

Carr, B. (2015). Higher dimensions of space and time and their implications for psi. In E. C. May & S. B. Marwaha (Eds.), *Extrasensory perception: Support, scepticism, and science. Vol. 2: Theories of psi* (pp.21-61). Santa Barbara, CA: Praeger.

Carter, C. (2010). "Heads I lose, tails you win", or, how Richard Wiseman nullifies positive results, and what to do about it: A response to Wiseman's (2010) critique of parapsychology. *Journal of the Society for Psychical Research, 74,* 156-167.

Chen, S. C. (1937). The leaders and followers among the ants in nest-building. *Physiological Zoolo*gy*, 10*(4), 437-455.

Claridge, G. (1997). *Schizotypy: Implications for illness and health.* Oxford: Oxford University Press.

Claridge, G. (2001). Spiritual experience: Healthy psychoticism? In I. Clarke (Ed.) *Psychosis and spirituality: Exploring the new frontier*, (pp. 90-106). Philadelphia, PA: Whurr Publishers.

Claridge, G., & Beech, T. (1995). Fully and quasi-dimensional constructions of schizotypy. In A. Raine, T. Lencz, & S. A. Mednick (Eds.), *Schizotypal personality, (*pp. 192-216). New York, NY: Cambridge University Press.

Claridge, G., & Davis, C. (2003). Personality and psychological disorders. London: Arnold.

Clarke, C. (2004). Quantum mechanics, consciousness and the self. In D. Lorimer (Eds.), *Science, Consciousness and Ultimate Reality*, Exeter: Imprint Academic.

Clouston, T. S. (1892). *Clinical lectures on mental diseases*. London: Churchill.

Colwell, J., S. Schroder., and D. Sladen. (2000). The ability to detect unseen staring: A literature review and empirical tests. *British Journal of Psychology, 91*, 71-85.

Conty, L., Gimmig, D., Belletier, C., George, N., & Huguet, P. (2010). The cost of being watched: Stroop interference increases under concomitant eye contact. *Cognition*, *115*(1), 33-139.

Cook, C. M., & Persinger, M. A. (1997). Experimental induction of the sensed presence in normal subjects and an exceptional subject. *Perceptual and Motor Skills, 85*, 683-693.

Coover, J. E. (1913). The feeling of being stared at. *American Journal of Psychology, 24,*570-575.

Corbett, J. (1986), *Jim Corbett's India.* Oxford: Oxford University Press.

Cottrell, N. B., Wack, D. L., Sekerak, G. J., & Rittle, R. H. (1968). Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology, 9*, 245-250.

Cottrell, J. E., Winer, G. A., & Smith, M. C. (1996). Beliefs of children and adults about feeling stares of unseen others. *Developmental Psychology, 3*(2), 50-61.

Dagnall, N., Parker, A., & Munley, G. (2007). Paranormal belief and reasoning. *Personality and Individual Differences, 43*, 1406-1415.

Dagnall, N., Denovan, A., Drinkwater, K., Parker, A.,& Clough, P. (2016). Toward a better understanding of the relationship between belief in the paranormal and statistical bias: The potential role of schizotypy. *Frontiers in Psychology, 7,* 1045.

Dahl, D. W., Manchanda, R. V., & Argo, J. J. (2001). Embarrassment in consumer purchase: The roles of social presence and purchase familiarity. *Journal of Consumer Research, 28*(3), 473-481.

Danckert, J., & Merrifield, C. (2018). Boredom, sustained attention and the default mode network. Experimental Brain Research, 236, 2507-2518.

Dashiell, J. F. (1935). Experimental studies of the influence of social situations on the behaviour of individual human adults. In C. Murchison (Ed.), *A handbook of social psychology* (pp. 1097-1158). Worcester, MA: Clark University Press.

Degerman, A., Rinne, T., Pekkola, J., Autti, T., Jääskeläinen, I. P., Sams, M., & Alho, K. (2007). Human brain activity associated with audiovisual perception and attention. *Neuroimage, 34*, 1683-1691.

Delbrück, M. (1986). *Mind from matter?* Palo Alto, CA: Blackwell Scientific Publications.

Deming, D. (2016). Do extraordinary claims require extraordinary evidence? *Philosophia, 44,* 1319-1331.

Dewhurst, K., & Beard, A. W. (1970). Sudden religious conversion in temporal lobe epilepsy. *British Journal of Psychiatry, 117*, 497-507.

Ducrocq, E., Wilson, M., Smith, T., & Derakshan, N. (2017). Adaptive working memory training reduces the negative impact of anxiety on competitive motor performance. *Journal of Sport and Exercise Psychology*, *39*(6), 412-422.

Eckbald, M., & Chapman, L. (1983). Magical ideation as an indicator of schizotypy. *Journal of Consulting and Clinical Psychology, 51*, 215-225.

Edge, H. L., Morris, R. L., Palmer, J., & Rush, J. H. (1986). *Foundations of parapsychology*. London: Routledge & Kegan Paul.

Eger, E., Ashburner, J., Haynes, J. D., Dolan, R. J., & Rees, G. (2008). FMRI activity patterns in human LOC carry information about object exemplars within category. *Journal of Cognitive Neuroscience, 20,* 356-370.

Ekdahl, A. G. (1929). The effect of attitude on free word association time. *Genetic, Social, and General Psychology Monographs, 5*, 253 - 338.

Esmark, C., Noble, S., & Breazeale, M. (2017). I'll be watching you: Shoppers' reactions to perceptions of being watched by employees. *Journal of Retailing*, *93*(3), 336-349.

Eysenck, H. (1995). *Genius: The natural history of creativity*. Cambridge: Cambridge University Press.

Falk, R. (1989). Judgment of coincidences: Mine vs. yours. *American Journal of Psychology, 102,* 477-493.

Fallon, J. H. (2015). Foreword. In E. C. May & S. B. Marwaha (Eds.), *Extrasensory perception: Support, scepticism, and science. Vol. 1: History, controversy, and research* (pp. ix–xv). Santa Barbara, CA: Praeger.

Farah, M. J., Weisberg, L. L., Monheit, M., & Peronnet, F. (1989). Brain activity underlying mental imagery: Event-related potentials during mental image generation. *Journal of Cognitive Neuroscience, 1*(4), 302-316.

Farnworth, B., Innes, J., Davy, M., Little, L., Cave, V., & Waas, J. (2020). Antipredator responses of ship rats to visual stimuli: Combining unimodal predation cues generates risk avoidance. *Animal Behaviour*, *168*, 149-157.

Farshadkhah, S., Van Slyke, C., & Fuller, B. (2021). Onlooker effect and affective responses in information security violation mitigation. *Computers & Security*, *100*, 1-16.

Fosse, R., Stickgold, R., & Hobson, A. (2004). Thinking and hallucinating: Reciprocal changes in sleep. *Psychophysiology, 41*, 298-305.

Fredrickson, B. (2002). Positive emotions. In C. Snyder & S. Lopez (Eds.), *Handbook of positive psychology,* (pp.120-134). New York, NY: Oxford University Press.

Freedman, A. (2005). The Sense of Being Glared At: What Is It Like to be a Heretic? *Journal of Consciousness Studies, 12*, 4-9.

French, C. (2001). Weird science at Goldsmiths. *Sceptic, 14*, 7-8.

Friday, R., & Luke, D. (2014, September). *I hear dead people: Individual differences in the perception of anomalous voices in ambiguous electronic audio recordings.* Paper presented to the International Conference of the Society for Psychical Research, University of Swansea, 6th September.

Friday, R., & Luke, D. (2014, September). *Now see hear! Detecting being watched or listened to via extrasensory means.* Paper presented to the International Conference of the Society for Psychical Research, University of York, 4th - 6th September.

Friedrich, J. (1993). Primary error detection and minimization (PEDMIN) strategies in social cognition: A reinterpretation of confirmation bias phenomena. *Psychological Review, 100*(2), 298-319.

Frith, C. (1979). Consciousness, information processing and schizophrenia. *British Journal of Psychiatry, 134*, 225-235.

Frith, C., & Raymond, J. D. (1997). Brain mechanisms associated with top-down processes in perception. *Philosophical Transactions of the Royal Society B, 352*, 1221-1230.

Gallagher, C., Kumar, V. K., & Pekala, R. J. (1994). The anomalous experiences inventory reliability and validity. *Journal of Parapsychology, 58*, 402-428.

Geisler, W. S., & Albrecht, D. G. (1997). Visual cortex neurons in monkeys and cats: Detection, discrimination, and identification. *Visual Neuroscience, 14*, 897-919.

Gerlach, A. L., Wilhelm, F. H, & Roth, W. T. (2003). Embarrassment and social phobia: The role of parasympathetic activation. *Journal of Anxiety Disorders, 17*(2), 197-210.

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences, 10*, 278-285.

Gianotti, L., Mohr, C., Pizzagalli, D., Lehmann, D., & Brugger, P. (2001). Associative processing and paranormal belief. *Psychiatry and Cognitive Neurosciences, 55*, 595-603.

Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on "Estimating the reproducibility of psychological science." *Science, 351*, 1037-1038.

Glicksohn, J. (1990). Belief in the paranormal and subjective paranormal experience. *Personality and Individual Differences, 11*, 675-683.

Gloor, P., Olivier, A., Quesney, L. F., Andermann, F., & Horowitz, S. (1982). The role of the limbic system in experiential phenomena of temporal lobe epilepsy. *Annals of Neurology, 12*,129-144.

Gnam, J. P., Loeffler, S. N., Haertel, S., Engel, F., Hey, S., Boes, K., Woll, A., & Strahler, J. (2019). On the relationship between physical activity, physical fitness, and stress reactivity to a real-life mental stressor. *International Journal of Stress Management*, *26*(4), 344-355.

Goddard, G. V., McIntyre, D. C., & Leech, C. K. (1969). A permanent change in brain function resulting from daily electrical stimulation. *Experimental Psychology, 25,* 295-330.

Goodale, M. A., & Milner, A. D. (2004). *Sight unseen*: *An exploration of conscious and unconscious vision.* New York: Oxford University Press.

Goodman, J., McClay, M., & Dunsmoor, J. (2020). Threat-induced modulation of hippocampal and striatal memory systems during navigation of a virtual environment. *Neurobiology of Learning and Memor*y, *168*:107160.

Gordon, J., & King, N. (2002). Children's night-time fears: an overview. *Counselling Psychology Quarterly*, *15*(2), 121-132.

Goulding, A. (2004). Schizotypy models in relation to subjective health and paranormal belief and experiences. *Personality and Individual Differences, 37*, 157-167.

Goulding, A. (2005). Healthy schizotypy in a population of paranormal believers and experients. *Personality and Individual Differences, 38*, 1069-1083.

Gray, J. (2004).*Consciousness: Creeping up on the hard problem.* Oxford: Oxford University Press.

Gray, N., Fernandez, M., Williams, J., Ruddle, R., & Snowden R. (2002). Which schizotypal dimensions abolish latent inhibition? *British Journal of Clinical Psychology*, *41*, 271-284.

Greenhouse, J. B. (1991). Comment: Parapsychology - On the margins of science. *Statistical Science, 6,* 386-389.

Griffin, M. (2001). The phenomenology of the alone condition: more evidence for the role of aloneness in social facilitation. *Journal of Psychology, 135*, 125.

Griffin, M., & Kent, M. V. (1998). The phenomenology of the alone condition. The role of aloneness in social facilitation. *The Journal of Social Psychology, 138*,667-669.

Groth-Marnat, G., & Pegden, J. A. (1998). Personality correlates of paranormal belief: Locus of control and sensation seeking. *Social Behaviour and Personality, 26,* 291-296.

Guerin, B. (1993). *Social facilitation*. Cambridge: Cambridge University Press.

Gupta, M., & Agrawal, A. (2012). A comprehensive review on systematic and meta-analysis methods. *International Journal of Pharmacology & Life Sciences, 3,* 1470-1474.

Hacking, I. (1988). Telepathy: Origins of randomization in experimental design. *Isis*, *79,* 427-451.

Haidich, A. B. (2010). Meta-analysis in medical research. *Hyppokratia*, *14,* 29-37.

Hair, J., Anderson, R., Tatham, R., & Black, W. (1998). *Multivariate data analysis*. Saddle River, NJ: Prentice-Hall International. Heppner, P.P., Kivlighan, D., and Wampold, B. (2008) *Research Design in Counselling*. Belmont, CA: Thomson Brooks/Cole.

Harris, M. J., & Rosenthal, R. (1988). *Interpersonal expectancy effects and human performance*. Washington, DC: National Academy Press.

Hergovich, A., Schott, R., & Arendasy, M. (2008). On the relationship between paranormal belief and schizotypy among adolescents. *Personality And Individual Differences*, *45*(2), 119-125.

Hergovich, A., Willinger, U., & Arendasy, M. (2005). Paranormal belief, schizotypy, and body mass index. *Perceptual & Motor Skills*, *100*(3), 883-891.

Hergovich, A., Schott, R., & Burger, C. (2010). Biased evaluation of abstracts depending on topic and conclusion: Further evidence of a confirmation bias within scientific psychology. *Current Psychology, 29*, 188-209.

Hershenson, M. (1998). *Visual space perception*. Cambridge: Mass.: MIT Press.

Hesslinger, V., Carbon, C., & Hecht, H. (2017). The Sense of Being Watched Is Modulated by Arousal and Duration of the Perceptual Episode. *I-Perception*, *8*(6), 1-11.

Hofstadter, D. (2011, January 7). A cut off for craziness. *The New York Times*. Retrieved from https://www.nytimes.com/roomfordebate/2011/01/06/the-esp-study-when-science-goes-psychic/a-cutoff-for-craziness.

Holt, N. (2007). Creativity, states of consciousness and anomalous cognition: The role of epistemological flexibility in the creative process. Unpublished Doctoral thesis, University of Northampton.

Holt, N., Simmonds-Moore, C., & Moore, S. (2008). Benign schizotypy: Investigating differences between clusters of schizotype on paranormal belief, creativity, intelligence and mental health. Proceedings of Presented Papers: The Parapsychological Association 51st Annual Convention, 82-96.

Honorton, C. (1977). Psi and attentional internal states. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp. 435-472). Jefferson, NC: Mc-Farland.

Horowitz, M. J., & Adams, J. E. (1970). Hallucinations on brain stimulation: Evidence for revision of the Penfield hypothesis. New York, NY: Plenum.

Horrey, W. J., Lesch, M. F., Garabet, A., Simmons, L., & Maikala, R. (2017). Distraction and task engagement: How interesting and boring information impact driving performance and subjective and physiological responses. *Applied Ergonomics, 58*, 342-348.

Hövelmann, G. H. (2015). On the usefulness of parapsychology for science at large. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 389-398). Jefferson, NC: McFarland.

Huguet, P., Galvaing, M. P., Monteil, J. M., & Dumas, F. (1999). Social presence effects in the Stroop task: Further evidence for an attentional view of social facilitation. *Journal of Personality and Social Psychology*, *77*(5), 1011-1025.

Hyman, R. (1994). Anomaly or artefact? Comments on Bem and Honorton. *Psychological Bulletin, 115,* 19-24.

Hyman, R., & Honorton, C. (1986). A joint communique´: The psi ganzfeld controversy. *Journal of Parapsychology, 50,* 350-364.

Irwin, H. (1991). Reasoning skills of paranormal believers. *Journal of Parapsychology, 55*, 281-300.

Irwin, H. J. (1997). An empirically derived typology of paranormal believers. *European Journal of Parapsychology, 13*, 1-14.

Irwin, H. J. (1999). *An introduction to parapsychology* (3rd ed.). Jefferson, North Carolina and London: McFarland and Co.

Irwin, H. J. (2009). *The psychology of paranormal belief: A researcher's handbook.* Hatfield: University of Hertfordshire Press.

Irwin, H. J., & Green, M. (1999). Schizotypal processes and belief in the paranormal: A multidimensional study. *European Journal of Parapsychology, 14,* 1-15.

Isen, A. (1999). On the relationship between affect and creative problem solving. In S. Russ (Ed.), *Affect, creative experience and psychological adjustment*, (pp. 3-17). Philadelphia, PA: Brunner/Mazel.

Jackson J. On cases of epileptic seizures with an auditory warning. *Lancet, 1876*, 386-387.

Jackson, M. (1997). Benign schizotypy? The case of spiritual experience. In G. Claridge (Eds.), Schizotypy: Implications for illness and health, (pp. 227-250). New York: Oxford University Press.

Jamieson, J., Black, A., Pelaia, L., & Reis, H. (2020). The impact of mathematics anxiety on stress appraisals, neuroendocrine responses, and academic performance in a community college sample. *Journal of Educational Psychology*. doi: 10.1037/edu0000636.

Jentsch, V., & Wolf, O. (2020). The impact of emotion regulation on cardiovascular, neuroendocrine and psychological stress responses. *Biological Psychology*, *154*, 107893. doi: 10.1016/j.biopsycho.2020.107893

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430-454.

Kapur, S. (2009). Schizophrenia. *Lancet, 374,* 635-645.

Kennedy, J. E. (2005). Personality and motivations to believe, misbelieve, and disbelieve in paranormal phenomena. *Journal of Parapsychology, 69*, 263-292.

Kennedy, J. E., & Kanthamani, H. (1995). An exploratory study of the effects of paranormal and spiritual experiences on peoples' lives and well-being. *The Journal of the American Society for Psychical Research, 89*, 249-265.

Kennedy, J., & Kanthamani, H. (1995). Association between anomalous experiences and artistic creativity and spirituality. *Journal of the American Society for Psychical Research, 89*, 333-341.

Kennedy, J., Kanthamani, H., & Palmer, J. (1994). Psychic and spiritual experiences, health and well-being, and meaning in life. *Journal of Parapsychology, 58*, 353-383.

Kille, L. W., & Maximino, M. (2014). The effect of CCTV on public safety: Research roundup. *Journalists Resource*. Harvard Kennedy School's Shorenstein Centre and the Carnegie-Knight Initiative. Retrieved 05-06-2015.

Kim, Y. J., Grabowecky, M., & Suzuki S. (2006). Stochastic resonance in binocular rivalry. *Vision Research, 46,* 392-406.

Klauer, K. C., Herfordt, J., & Voss, A. (2008). Social presence effects on the stroop-task: Boundary conditions and an alternative account. *Journal of Experimental Social* Psychology, 44, 469-476.

Kohr, R. L. (1980). A survey of psi experiences among members of a special population. *The Journal of the American Society for Psychical Research, 74,* 395-411.

Lange, R., & Houran, J. (1997). Context induced paranormal experiences: Support for Houran and Lange's model of haunting phenomena. *Perceptual and Motor Skills, 84,* 1455-1458.

Lange, R., & Houran, J. (1998). Delusions of the paranormal: A haunting question of perception. *Journal of Nervous and Mental Disease, 186,* 637-645.

Lange, R., & Houran, J. (1999). The role of fear in delusions of the paranormal. *Journal of Nervous and Mental Disease, 187,* 159-166.

Lawrence, T. R. (1993). Gathering in the sheep and goats: A meta-analysis of forced-choice sheep-goat ESP studies, 1947-1993. *Proceedings of Presented Papers: The Parapsychological Association 36th Annual Convention,* 75-86.

Lawrence, T. R. (1995). How many factors of paranormal belief are there? A critique of the Paranormal Belief Scale. *The Journal of Parapsychology, 59,* 3-25.

Lawrence, T., Edwards, C., Barraclough, N., Church, S., & Hetherington, F. (1995). Modelling childhood causes of paranormal belief and experience: Childhood trauma and childhood fantasy. *Personality and Individual Differences, 19,* 209-215.

Laythe, B., & Owen, K. (2012). Paranormal belief and the strange case of haunt experiences: evidence of a neglected population. *The Journal of Parapsychology, 76,* 23-29.

Lange, R., Thalbourne, M. A., Houran, J., & Storm, L. (2000). The revised transliminality scale: Reliability and validity data from a Rasch top-down purification procedure. *Consciousness and Cognition, 9*(4), 591-617.

Lee, B., Hadjiosif, M., Wedlich, D., Barlow, L., Crichton, C., & Moorehead, C. (2002). Personality and Perception in Staring Detection and Experience. *The Parapsychological Association 45th Annual Convention Proceedings of Presented Papers*, 407-409.

Lehar, S. (2003). 'Gestalt isomorphism and the primacy of subjective conscious experience: A gestalt bubble model.' *Behavioural and Brain Sciences*, *26,* 375-444.

Lewontin, R. (1994). A rejoinder to William Wimsatt. In J. Chandler, A. I. Davidson, & H. D. Harootunian (Eds.), *Questions of evidence: Proof, practice, and persuasion across the disciplines* (pp. 504-09). Chicago, IL: University of Chicago Press.

Liang, L., Lu, T., & Wang, X. (2002). Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates, *Journal of Neurophysiology*. *87,* 2237-2261.

Lindberg, D.C. (1981). *Theories of vision from Al-Kindi to Kepler*. Chicago, IL: University of Chicago Press.

Lindeman, M., & Aarnio, K. (2006). Paranormal beliefs: Their dimensionality and correlates. *European Journal of Personality, 20,* 585-602.

Lindeman, M., Cederstrom, S., Simola, P., Simula, A., Ollikainen, S., & Riekki, T. (2008). Sentences with core knowledge violations increase the size of N400 among paranormal believers. *Cortex, 44*, 1307-1315.

Lobach, E., & Bierman, D. (2004). The invisible gaze: Three attempts to replicate Sheldrake's staring effects. Proceedings of the 47th PA Convention, 77-90.

Loughland, C., & Williams, L. (1997). A cluster analytic study of schizotypal trait dimensions. *Personality and Individual Differences, 23,* 877-883.

Luckoff, D. (2007). Visionary spiritual experiences. *Southern Medical Journal, 100*, 635-641.

Luke, D., Firth, B., & Pilkington, M. (2013, September*). The men (and women) who stare at (sheep and) goats: Beliefs, expectations, experiences, neurology and gender in haunt site vigils.* Paper presented to the International Conference of the Society for Psychical Research, University of Swansea, 6th September.

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health, 2*, 151-169.

Lynn, S., Neufeld, V., Green, J., Sandberg, D., & Rhue, J. (1996). Daydreaming, fantasy and psychopathology. In R. Kunzendorf, N. Spanos, & B. Wallace (Eds.), *Hypnosis and Imagination*, (pp. 67-98). New York: Baywood Publishing Company.

MacCoun, R. J. (1998). Biases in the interpretation and use of research results. *Annual Review of Psychology, 49,* 259-287.

Makarec, K., & Persinger, M. A. (1990). Electroencephalographic validation of a temporal lobe signs inventory. *Journal of Research in Personality, 24,* 323-337.

Marks, D.F. (2000). *The psychology of the psychic* (2nd Ed). Amherst, New York: Prometheus Books.

Marks, D., & Colwell, J. (2000). The psychic staring effect: An artefact of pseudo randomization. *Sceptical Inquirer,* 41-49.

Marks, D., & Kammann, R. (1980). *The psychology of the psychic*. Buffalo, New York: Prometheus Books.

Markus, H. (1978). That effect of mere presence on social facilitation: An unobtrusive test. *Journal of Experimental Social Psychology, 14*, 398-397.

Martinez, E. G., Ortega, L., Grabowecky, M., Mossbridge, J., & Suzuki, S. (2012). Interactive coding of visual spatial frequency and auditory amplitude-modulation rate. *Current Biology, 22*(5), 383-388.

Marwaha, S. B., & May, E. C. (2016). Precognition: The only form of psi? *Journal of Consciousness Studies, 23,* 76-100.

Mason, O., Claridge, G., & Jackson, M. (1995). New scales for measurement of schizotypy. *Personality and Individual Differences, 18*, 7-13.

Mason, O., Claridge, G., & Williams, L. (1997). Questionnaire measurement. In G. Claridge (Ed.), *Schizotypy: Implications for illness and health*, pp. 19-3). New York: Oxford University Press.

Mason, O., Linney, Y., & Claridge, G. (2005). Short scales for measuring schizotypy - The Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE). *Schizophrenia Research, 78*, 293-296.

Mathijsen, F. (2016). Anomalous experiences and schizotypy: which comes first? The Hermit Crab syndrome hypothesis. *Mental Health, Religion & Culture*, *19*(2), 113-123.

May, E. C., & Marwaha, S. B. (2015). *Extrasensory perception: Support, scepticism, and science* (Vols. 1-2). Santa Barbara, CA: Praeger.

McClenon, J. (1994). *Wondrous events: Foundations of religious belief*. Philadelphia: University of Pennsylvania Press.

McCreery, C., & Claridge, G. (1995a). A study of hallucination in normal subjects – 1. Self-report data. *Personality and Individual Differences, 21*, 739-747.

McCreery, C., & Claridge, G. (1995b). Out-of-the-body experiences and personality. *Journal of the Society for Psychical Research, 60,* 129-148.

McCreery, C., & Claridge, G. (2002). Healthy schizotypy: The case of out-of-body experiences. *Personality and Individual Differences, 32,* 141-154.

Mehrabian, A., Stefl, C., & Mullen, M. (1997). Emotional thinking in the adult: Individual differences in mysticism and globality-differentiation. *Imagination, Cognition and Personality, 16*, 325-355.

Metzinger, T. (2005). Out-of-body experiences as the origin of the concept of a soul. *Mind and Matter, 3*, 57-84.

Miller, G. A. (2006). Retrieved 04.05.2010, fromwordnetweb.princeton.edu/perl/webwn.

Millar, B. (2015). Quantum theory and parapsychology. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 165-180). Jefferson, NC: McFarland.

Milton, J., & Wiseman, R. (1999). Does Psi Exist? Lack of replication of an anomalous process of information transfer. *Psychological Bulletin, 125*, 387-391.

Mostajeran, F., Balci, M. B., Steinicke, F., Kühn, S., & Gallinat, J. (2020). The effects of virtual audience size on social anxiety during public speaking. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Atlanta, GA, USA*, *2020* (pp. 303-312).

Nagel, T. (2012). Mind and cosmos: Why the materialist neo-Darwinian conception of nature is almost certainly false. Oxford, United Kingdom: Oxford University Press.

Nakagawa, Y., Park, K., Ueda, H., Ono, H., & Miyake, H. (2019). Being watched over by a conversation robot may enhance safety in simulated driving. *Journal of Safety Research*, *71*, 207-218.

Neppe, V. M. (1983). Temporal lobe symptomatology in subjective paranormal experiments. *Journal of the American Society for Psychical Research*, *77,* 1-29.

Nettle, D. (2006). Schizotypy and mental health amongst poets, visual artists, and mathematicians. *Journal of Research in Personality, 40,* 876-890.

Nunnally, J. C. (1978). *Psychometric theory* (2^{nd}ed.). New York: McGraw-Hill.

Office of Technology Assessment. (1989). Report of a workshop on experimental parapsychology. *Journal of the American Society for Psychical Research, 83,* 317-339.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*.

O'Sullivan, G., McGuire, B., Roche, M., & Caes, L. (2020). Am I being watched? The role of researcher presence on toddlers' behaviour during 'everyday' pain experiences: a pilot study. *Psychology & Health*, *35*(9), 115-119.

Özgen, E., Payne, H. E., Sowden, P. T., & Schyns, P. G. (2006). Retinotopic sensitisation to spatial scale: Evidence for flexible spatial frequency processing in scene perception. *Vision Research, 46,* 1108-1119.

Palmer, J. (1979). A community mail survey of psychic experiences. *Journal of the American Society for Psychical Research*, *73*, 221-251.

Palmer, J. (1986). Terminological poverty in parapsychology: Two examples. In D.H. Weiner & D. Radin (Eds.), *Research in Parapsychology 1985*. Metuchen, NJ: Scarecrow Press.

Palmer, J., & Millar, B. (2015). Experimenter effects in parapsychology research. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 293-300). Jefferson, NC: McFarland.

Parsons, T. D., & Courtney, C., G. (2018). Interactions Between Threat and Executive Control in a Virtual Reality Stroop Task. *Transactions on Affective Computing, 9*(1), 66-67.

Pavlenko, D., & Akiyama, T. (2019). Why does stress aggravate itch? A possible role of the amygdala. *Experimental Dermatology*, *28*(12), 1439-1441.

Peng, A. W., Salles, F. T., Pan, B., & Ricci, A. J. (2011). Integrating the biophysical and molecular mechanisms of auditory hair cell mechanotransduction. *Nature Communications, 1*(2), 523-524.

Persinger, M.A. (1984). Propensity to report paranormal experiences is correlated with temporal lobe signs. *Perceptual and Motor Skills, 59*, 583-586

Persinger, M. A. (2001). The neuropsychiatry of paranormal experiences. *Journal of Neuropsychiatry and Clinical Neurosciences, 13*, 515-524.

Persinger, M.A., & Makarec, K. (1993). Complex partial epileptic signs as a continuum from normals to epileptics: Normative data and clinical populations. *Journal of Clinical Psychology, 49*, 33-45.

Persinger, M. A., Tiller, S. G., & Koren, S. A. (2000). Experimental simulation of a haunt experience and elicitation of paroxysmal electroencephalographic activity by trans cerebral complex magnetic fields: Induction of a synthetic ghost. *Perceptual Motor Skills, 90*, 659-674.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience, 9*, 148-158.

Peterson, D.M. (1978). Through the looking glass: An investigation of the faculty of extra sensory detection of being looked at. Unpublished MA thesis, University of Edinburgh.

Pfattheicher, S., & Keller, J. (2015). The watching eyes phenomenon: The role of a sense of being seen and public self-awareness. *European Journal Of Social Psychology*, *45*(5), 560-566.

Pfattheicher, S., Schindler, S., & Nockur, L. (2019). On the impact of Honesty-Humility and a cue of being watched on cheating behaviour. *Journal of Economic Psychology, 71,* 159-174.

Planck, Max K. (1950). *Scientific Autobiography and Other Papers*. New York: Philosophical library.

Platt, J. J., Yaksh, T., & Darby, C. L. (1967). Social facilitation of eating behaviour in armadillos. *Psychological Reports, 20*(2), 1136.

Poortman, J. J. (1959). The feeling of being stared at. *Journal of the Society for Psychical Research*, *40*, 4-12.

Posey, T. & Losch, M. (1983). Auditory hallucinations of hearing voices in 375 normal subjects. *Imagination, Cognition and Personality, 3*, 99-113.

Rabeyron, T. & Watt, C. (2010). Paranormal experiences, mental health and mental boundaries, and psi. *Personality and Individual Differences*, *48*(4), 487-492.

Rachman, S. (1977) The conditioning theory of fear acquisition: A critical examination. *Behaviour Research & Therapy, 15*, 375-387.

Radin, D. (2004a). Entangled minds. *Shift*, *5*, 10-14.

Radin, D. (2004b). The feeling of being stared at: An analysis and replication. *Journal of the Society for Psychical Research*, *68*, 245-252.

Radin, D., Michel, L., & Delorme, A. (2016). Psychophysical modulation of fringe visibility in distant double-slit optical system. *Physics Essays, 29,* 14-22.

Ramachandran, V. S. (1998). Phantoms in the Brain: Probing the Mysteries of the Human Mind. New York, NY: William Morrow.

Rawlings, D., Williams, B., Haslam, N., & Claridge, G. (2008). Taxometric analysis supports a dimensional latent structure for schizotypy. *Personality and Individual Differences, 44,* 1640-1651.

Renaud, P., & Blondin, J. (1997). The stress of Stroop performance: Physiological and emotional responses to colour-word interference, task pacing, and pacing speed. *International Journal of Psychophysiology, 27*(2), 87-97.

Rhine, J. B. (1934). *Extra-sensory perception*. Boston: Boston Society for Psychic Research.

Rice, T. (2004). Believe it or not: Religious and other paranormal beliefs in the United States. *Journal for the Scientific Study of Religion, 42*, 95-106.

Richards, A., French, C. C., Johnson, W., Naparstek, J., &Williams, J. (1992). Effects of mood manipulation and anxiety on performance of an emotional Stroop task. *British Journal of Psychology, 83*(4), 479-491.

Richards, A., Richards, L. C., & McGeeney, A. (2000). Anxiety-Related Stroop Interference in Adolescents. *The Journal of General Psychology, 127*(3), 327-333.

Ritchie, S. J., Wiseman, R., & French, C. C. (2012). Failing the future: Three unsuccessful attempts to replicate Bem's 'retroactive facilitation of recall' effect. *PLoS ONE, 7(3), e*33423.

Ridley, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology,18*(6), 643-662.

Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology, 7,* 331-363.

Ring, K. (1984). Heading toward omega: In search of the meaning of the near-death experience. New York, NY: William Morrow.

Robbins, B., D., & Parlavecchio, H. (2009). The unwanted exposure of the self: A phenomenological study of embarrassment. *The Humanistic Psychologist, 34*(4), 321 - 345.

Robertson, L. (2004). *Space, objects, minds, and brains*. New York: Psychology Press.

Roig, M., Bridges, R., Hackett Renner, C., & Jackson, C. (1998). Belief in the paranormal and its association with irrational thinking controlled for context effects. *Personality and Individual Differences, 24*, 229-236.

Roney-Dougal, S. M., Ryan, A., & Luke, D. (2014). The relationship between local geomagnetic activity and psychic awareness. *Journal of Parapsychology, 78(2)*, 235-254.

Rosenthal, R. (1976). *Experimenter effects in behavioural research.* New York, NY: John Wiley.

Rosenthal, R. (1990). Replication in behavioural research. *Journal of Social Behaviour and Personality, 5,* 1-30.

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science, 3*, 176-179.

Ross, C., & Joshi, S. (1992). Paranormal experiences in the general population. *The Journal of Nervous and Mental Disease, 180,* 357-361

Rouder, J. N., Morey, R. D., & Province, J. M. (2013). A Bayes factor meta-analysis of recent extrasensory perception experiments: Comment on Storm, Tressoldi, and Di Risio (2010). *Psychological Bulletin, 139,* 241-247.

Safina, C. (2015). *Beyond words: What animals think and feel*. New York, NY: Henry Holt.

Sarigiannidis, I., Grillon, C., Ernst, M., Roiser, J., & Robinson, O. (2020). Anxiety makes time pass quicker while fear has no effect. *Cognition, 197*, 104116.

Sarigiannidis, I., Kirk, P., Roiser, J., & Robinson, O. (2020). Does overloading cognitive resources mimic the impact of anxiety on temporal cognition? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(10), 1828-1835.

Schachter, S. (1959). *The psychology of affiliation*. Stanford, California: Stanford University Press.

Schiele, M., Bandelow, B., Baldwin, D., Pini, S., & Domschke, K. (2020). A neurobiological framework of separation anxiety and related phenotypes. *European Neuropsychopharmacology*, *33*, 45-57.

Schlitz, M. J., & Braud, W. G. (1997). Distant intentionality and healing: Assessing the evidence. *Alternative Therapies, 3*, 62-73.

Schlitz, M. J., & LaBerge, S. (1994). Autonomic detection of remote observation: Two conceptual replications. *Proceedings of Presented Papers: The 37th Annual Convention of the Parapsychological Association*, 352-360.

Schlitz, M. J., & LaBerge, S. (1997). Covert observation increases skin conductance in subjects unaware of when they are being observed: A replication. *The Journal of Parapsychology, 61*, 185-196.

Schmidt, S. (2001). Empirical test of the theory of morphic resonance - Do we feel if somebody is staring at us?. *Forschende Komplementärmedizin, 8*, 48-50.

Schmidt, S., Gnam, J., Kopf, M., Rathgeber, T., & Woll, A. (2020). The Influence of Cortisol, Flow, and Anxiety on Performance in E-Sports: A Field Study. *Biomed Research International*, 1-6.

Schmidt, S., & Walach, H. (2000). Electrodermal activity (EDA) - state-of-the-art measurement and techniques for parapsychological purposes. *The Journal of Parapsychology*, *64*, 139 - 163.

Schmidt, S., & Walach, H. (2009). The feeling of being stared at: A parapsychological classic with a facelift. *European Journal of Parapsychology, 24,*117-138.

Schmitt, B. H., Gilovich, T., Goore, N., & Joseph, L. (1986). Mere presence and social facilitation: One more time. *Journal of Experimental Social Psychology*, *22*, 242-248.

Schmidt, S., Schneider, R., Utts, J., & Walach, H. (2004). Distant intentionality and the feeling of being stared at: Two meta-analyses. *British Journal of Psychology, 95*, 235-247.

Schofield, K., & Claridge, G. (2007). Paranormal experiences and mental health: Schizotypy as an underlying factor. *Personality and Individual Differences, 43,* 1908-1916.

Schooler, J. W. (2014). Metascience could rescue the 'replication crisis', *Nature, 515*, 9.

Schumaker, J. K. (1987). Mental health, belief deficit compensation and paranormal beliefs. *Journal of Psychology: Interdisciplinary and Applied, 121*, 451-457.

Schwartz, J. L., Berthommier, F., & Savariaux, C. (2004). Seeing to hear better: Evidence for early audio-visual interactions in speech identification. *Cognition*, *93,* 69-78.

Schwartz, G. E. R., & Russek, L. G. S. (1999). Registration of actual and intended eye gaze: Correlation with spiritual beliefs and experiences. *Journal of Scientific Exploration, 13*, 213-229.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270,* 303-304.

Sheehan, D. P. (2011). Quantum retro causation: Theory and experiment (AIP Conference Proceedings, Vol. 1408). Melville, NY: American Institute of Physics.

Sheehan, D. P. (2015). Remembrance of things future: A case for retro causation and precognition. In E. C. May & S. B. Marwaha (Eds.), *Extrasensory perception: Support, scepticism, and science* (pp. 85–109). Santa Barbara, CA: Praeger.

Sheils, D. (1978). A cross-cultural study of beliefs in out-of-the-body experiences, waking and sleeping. *Journal of the Society for Psychical Research, 49,* 697-741.

Sheldrake, R. (1981), A new science of life: The hypothesis of formative causation. London: Blond and Briggs.

Sheldrake, R. (1988). The presence of the past: Morphic resonance and the habits of nature. London: Collins.

Sheldrake, R. (1994). Seven experiments that could change the world: A do-it-yourself-guide to revolutionary science. London: Fourth Estate.

Sheldrake, R. (1996). *An experiment with birds*. Rome: Zerynthia.

Sheldrake, R. (1998). The sense of being stared at: Experiments in schools. *Journal of the Society for Psychical Research, 62,* 311-23.

Sheldrake, R. (1999). The 'sense of being stared at' confirmed by simple experiments. *Biology Forum, 92,* 53-76.

Sheldrake, R. (1999a). *Dogs that know when their owners are coming home*. New York: Crown Publishers.

Sheldrake, R. (1999b). Commentary on a paper by Wiseman, Smith, and Milton on the 'psychic pet' phenomenon. *Journal of the Society for Psychical Research, 63* (857), 306-311.

Sheldrake, R. (2000). A dog that seems to know when his owner is coming home: Videotape experiments and observations. *Journal of Scientific Exploration, 14*(2), 233-255.

Sheldrake, R. (2000). The 'sense of being stared at' does not depend on known sensory clues. *Biology Forum, 93,* 209-224.

Sheldrake, R. (2001a). Experiments on the sense of being stared at: The elimination of possible artefacts. *Journal of the Society for Psychical Research, 65,* 122-137.

Sheldrake, R. (2001b). Research on the sense of being stared at. *Sceptical Inquirer,* 58-61.

Sheldrake, R. (2002). The sense of being stared at: An experiment at Holma College. *Gränsoverskridaren, 10,* 21-23.

Sheldrake, R. (2003). The Sense of Being Stared At and other aspects of the extended mind. St Ives: Random House.

Sheldrake, R. (2003a). The sense of being stared at and other aspects of the extended mind. New York, NY: Crown Publishers.

Sheldrake, R. (2003b). The need for open-minded scepticism: A reply to David Marks. *The Skeptic*, *16*, 8-13.

Sheldrake, R. (2005a). The non-visual detection of staring. *Journal of Consciousness Studies, 12*, 117-126.

Sheldrake, R. (2005b). The sense of being stared at part 1: Is it real or illusory? *Journal of Consciousness Studies, 12*, 10-31.

Sheldrake, R. (2015). Psi in everyday life: Nonhuman and human. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 350-363). Jefferson, NC: McFarland.

Shermer, M. (2005). Rupert's resonance: The theory of "morphic resonance" posits that people have a sense of when they are being stared at. What does the research show? *Scientific American*, 293-295.

Shulman, G. L., Sullivan, M. A., Gish, K., & Sakoda, W. J. (1986). The role of spatial-frequency channels in the perception of local and global structure, *Perception, 15,* 259-273.

Shumaker, J. F. (2001). Mental health, belief deficit compensation, and paranormal belief. *Journal of Personality, 121,* 451-457.

Sidgwick, H. (1894). Report on a census on hallucinations. *Proceedings of the Society for Psychical Research, 26,* 259-394.

Silverthorn, D. U., Garrison, C. W., Johnson, B. R., & Silverthorn, A. C. (2009). *Human physiology: An integrated approach* (4th ed.). San Francisco: Pearson/Benjamin Cummings.

Simmonds, C. (2003). Investigating schizotypy as an anomaly-prone personality. Unpublished PhD thesis. University of Leicester/University College Northampton.

Simmonds-Moore, C. (2010). Sleep patterns, personality, and subjective anomalous experiences. *Imagination,Cognition and Personality, 29*(1), 71-86.

Simmonds-Moore, C., & Holt, N. (2007). Trait, state and psi: An exploration of the interaction between individual differences, state preference and psi performance in the Ganzfeld and a waking ESP control. *Journal of the Society for Psychical Research, 71*, 197-215.

Smith, C. L., Johnson, J. L., & Hathaway, W. L. (2009). Personality contributions to beliefs in paranormal phenomena. *Individual Differences Research, 7,* 65-96.

Sommer, A. (2013). Crossing the boundaries of mind and body: Psychical research and the origins of modern psychology. University College, London, London, United Kingdom.

Sowden, P. T, & Schyns, P. G. (2006). Channel surfing in the visual brain. *Trends in Cognitive Science, 10,* 538-545.

Spencer, B. (1995). Correlations, sample size, and practical significance: A comparison of psychological and medical investigations. *The Journal of Psychology, 129,* 469-475.

Spratling, W. P. (1904). *Epilepsy and its treatment.* Philadelphia: W. B. Saunders.

Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience, 27,* 279-306.

Stanford, R. G. (1974). An experimentally testable model for spontaneous psi events. *Journal* of the American Society for Psychical Research, 68, 34-57.

Stanford, R. G. (1982). An experimentally testable model for spontaneous extrasensory events. In Grattan-Guinness (Ed.), *Psychical Research: A guide to its history, principles and practices* (pp. 195 - 205). Wellinborough: Aquarian Press.

Stanford, R. G. (1990). An experimentally testable model for spontaneous psi events: A review of related evidence and concepts from parapsychology and othersciences. In S. Krippner (Ed.), *Advances in Parapsychological Research* (pp. 54-161). Jefferson, NC: McFarland.

Stanford, R. G. (2015). Psychological concepts of psi function: A review and constructive critique. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 94-109). Jefferson, NC: McFarland.

Stapp, H. P. (2017). *Quantum theory and free will*. New York, NY: Springer.

Steinkamp, F. (2005). Forced-choice ESP experiments: Their past and their future. In M. A. Thalbourne & L. Storm (Eds.), *Parapsychology in the 21st century: Essays on the future of psychical research* (pp. 124-163). Jefferson, NC: McFarland.

Storm, L., Tressoldi, P. E., & Di Risio, L. (2010a). A meta-analysis with nothing to hide: Reply to Hyman (2010). *Psychological Bulletin, 136*, 491-494.

Storm, L., Tressoldi, P. E., & Di Risio, L. (2010b). Meta-analysis of free-response studies, 1992-2008: Assessing the noise reduction model in parapsychology. *Psychological Bulletin, 136,* 471-485.

Strauss, B. (2002). Social facilitation in motor tasks: A review of research and theory. *Psychology of Sport and Exercise. 3*(3), 237-256.

Suhr, J., & Spitznagel, M. (2001). Factor versus cluster models of schizotypal traits. A comparison of unselected and highly schizotypal samples. *Schizophrenia Research, 52*, 231-239.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics (5th ed.).* Boston: Allyn and Bacon.

Targ, E., Schlitz, M., & Irwin, H. (2000). Psi-related experiences. In E. Cardeña, S. Lynn, & Krippner, S. (Eds.), *Varieties of anomalous experience: Examining the scientific evidence*, (pp. 219-252). Washington, DC: American Psychological Association.

Thalbourne, M. A. (2001). Measures of the sheep-goat variable, transliminality, and their correlates. *Psychological Reports*, *88*, 339-350.

Thalbourne, M. A. (2003). *A Glossary of Terms used in Parapsychology* (2nd ed.). London: William Heinemann.

Thalbourne, M. A. (2004). Glossary [of terms used in parapsychology, 2003. Puente Publications: Charlottesville, VA.]. *Journal of Parapsychology, 68*(1), 203-208.

Thalbourne, M. A., & Delin, P. S. (1993). A new instrument for measuring the sheep-goat variable: Its psychometric properties and factor structure. *Journal of the Society for Psychical Research*, *59,* 172-86.

Thalbourne, M., & Delin, P. (1994). A common thread underlying belief in the paranormal, creative personality, mystical experience and psychopathology. *Journal of Parapsychology*, *58*, 3-38.

Thalbourne, M. A., & Evans, L. (1992). Attitudes and beliefs about, and reactions to, staring and being stared at. *Journal for the Society for Psychical Research*, *58,* 380-385.

Thalbourne, M., & French, C. (1995). Paranormal belief, manic depressiveness and magical ideation: A replication. *Personality and Individual Differences, 18*, 291-292.

Thouless, R. H. (1942). Experiments on paranormal guessing. *British Journal of Psychology,33*, 15-27.

Tillmann, J. (2009). Fear of the dark: Night-time roosting and anti-predation behaviour in the grey partridge. *Behaviour, 146*(7), 999-1023

Titchener, E.B. (1898). The feeling of being stared at. *Science New Series*, *8*, 895-897.

Tobacyk, J. (1988). *A Revised Paranormal Belief Scale*. Unpublished manuscript. Louisiana Tech University, Ruston, LA.

Tobacyk, J. (2004). The Revised Paranormal Belief Scale. *International Journal of Transpersonal Studies, 23*, 94-98.

Tobacyk, J., & Milford, G. (1983). Belief in paranormal phenomena: Assessment instrument development and implications for personality functioning. *Journal of Personality and Social Psychology, 44*, 648-655.

Travis, L. E. (1925). The effect of a small audience upon eye-hand coordination. *The Journal of Abnormal and Social Psychology, 20,* 142-146.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *American Journal of Psychology, 9*, 507-533.

Tsai, C., Yeh, S., Huang, Y., Wu, Z., Cui, J., & Zheng, L. (2018). The Effect of Augmented Reality and Virtual Reality on Inducing Anxiety for Exposure Therapy: A Comparison Using Heart Rate Variability. *Journal of Healthcare Engineering, 2018,* 1-8.

UITP Activity report 2014 & 2015. Published on Jan 20, 2016. https://issuu.com/uitp3/docs/activity_report_2014_2015

Utts, J. (1991). Replication and meta-analysis in parapsychology. *Statistical Science, 6*(4), 363-403.

Uziel, L. (2007). Individual differences in the social facilitation effect: A review andmeta-analysis. *Journal of Research in Personality, 41*, 579-601.

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioural Statistics, 35,* 215-247.

Van Ee, R., Van Boxtel, J. J. A., Parker, A. L., & Alais, D. (2009). Multisensory congruency as a mechanism for attentional control over perceptual selection. *Journal of Neuroscience, 29,* 11641-11649.

Varlinskaya, E., Johnson, J., Przybysz, K., Deak, T., & Diaz, M. (2020). Adolescent forced swim stress increases social anxiety-like behaviours and alters kappa opioid receptor function in the basolateral amygdala of male rats. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *98*, 109812. doi: 10.1016/j.pnpbp.2019.109812

Velmans, M. (1990). Consciousness, brain and the physical world. *Philosophical Psychology*, *3*, 77-99.

Velmans, M. (2000). *Understanding consciousness*. London, United Kingdom: Routledge.

Velmans, M. (2007)."An Epistemology for the Study of Consciousness". In M. Velmans & S. Schneider (Eds.), *the Blackwell Companion to Consciousness* (pp. 711-725). New York, NY: Wiley-Blackwell.

Velmans, M. (2009). *Understanding consciousness*. 2nd ed. London: Routledge.

Velmans, M. (2018). How to investigate perceptual projection: A commentary on Pereira Jr., “The projective theory of consciousness: From neuroscience to philosophical psychology”. *Trans/Form/Ação*, *41*, 233-242.

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem. *Journal of Personality and Social Psychology*, *100,* 426-432.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly. Journal of Experimental Psychology, 12*(3), 129-140.

Watt, C., & Nagtegaal, M. (2004). Reporting of blind methods: An interdisciplinary survey. *Journal of the Society for Psychical Research, 68,*105-114.

Watt, C., & Tierney, I. (2014). Psi-related anomalous experiences. In E. Cardeña, S. J. Lynn, & S. Krippner (Eds.), *Varieties of anomalous experience: Examining the scientific evidence* (2nd ed.; pp. 241–272). Washington, DC: American Psychological Association.

Weber, B., Wellmer, J., Reuben, M., Mormann, F., Weis, S., Urbach, H., Ruhlmann, J., Elger, C. E., & Fernandez, G. (2006). Left hippocampal pathology is associated with atypical language lateralization in patients with focal epilepsy. *Brain*, *129*, 346-351.

Weinberg, R. (2011). *Foundations of sport and exercise psychology*. Champaign, IL: Human Kinetics.

West, L. (1962). *Hallucinations*. New York and London: Grune & Stratton.

Whiteside, S., Brennan, E., Biggs, B., Vickers, K., Hathaway, J., & Seifert, S. (2020). The feasibility of verbal and virtual reality exposure for youth with academic performance worry. *Journal of Anxiety Disorders*, *76*, 102298.

Williams, L. (1983). Minimal cue perception of the regard of others: The feeling of being stared at. *Journal of Parapsychology*, *47,* 59-60.

Williams, L. (1994). The multidimensional nature of schizotypal traits, a cluster analytic study. *Personality and Individual Differences*, *16*, 103-112.

Winer, G. A., Cottrell, J. E., Gregg, V. A., Fournier, J. S., & Bica, L.A. (2002). Fundamentally misunderstanding visual perception: Adults' beliefs in visual emissions. *American Psychologist*, *57*, 417-424.

Wiseman, R. (2010). Heads I win, tails you lose: how parapsychologists nullify null results. *Skeptical Inquirer*, *34*, 36-39.

Wiseman, R., & Greening, E. (2005). 'It's still bending'. Verbal suggestion and alleged psychokinetic ability. *British Journal of Psychology, 96,* 115-127.

Wiseman, R., & Schlitz, M. (1997). Experimenter effects and the remote detection of staring. *Journal of Parapsychology, 61*, 197-207.

Wiseman, R., & Schlitz, M. (1998). Experimenter effects and the remote detection of staring. *Journal of Parapsychology, 61*(3), 197-208.

Wiseman, R., & Schlitz, M. (1999). Replication of experimenter effect and the remote detection of staring. *The Journal of Parapsychology*, *63*(3), 232.

Wiseman, R., & M. D. Smith. (1994). *A further look at the detection of unseen gaze*. Proceedings of Presented Papers 37th Annual Convention. D. J. Bierman (Ed.), 465-478. Parapsychological Association, Fairhaven, Mass.

Wiseman, R., Smith, D., Freedman, T., Wasserman, T., & Hurst, C. (1995). Two further experiments concerning the remote detection of unseen gaze. Proceedings of Presented Papers 38th Annual Convention. Ed. by D. J. Bierman, 480-492. Parapsychological Association, Fairhaven, Mass.

Wiseman, R., Watt, C., Stevens, P., Greening, E., & O'Keeffe, C. (2003). An investigation into alleged "hauntings." *British Journal of Psychology*, *94*, 195-211.

Yau, J. M., Olenczak, J. B., Dammann, J. F., & Bensmaia, S. J. (2009). Temporal frequency channels are linked across audition and touch. *Current Biology, 19*, 561-566.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology*. *18*(5)*:* 459-482.

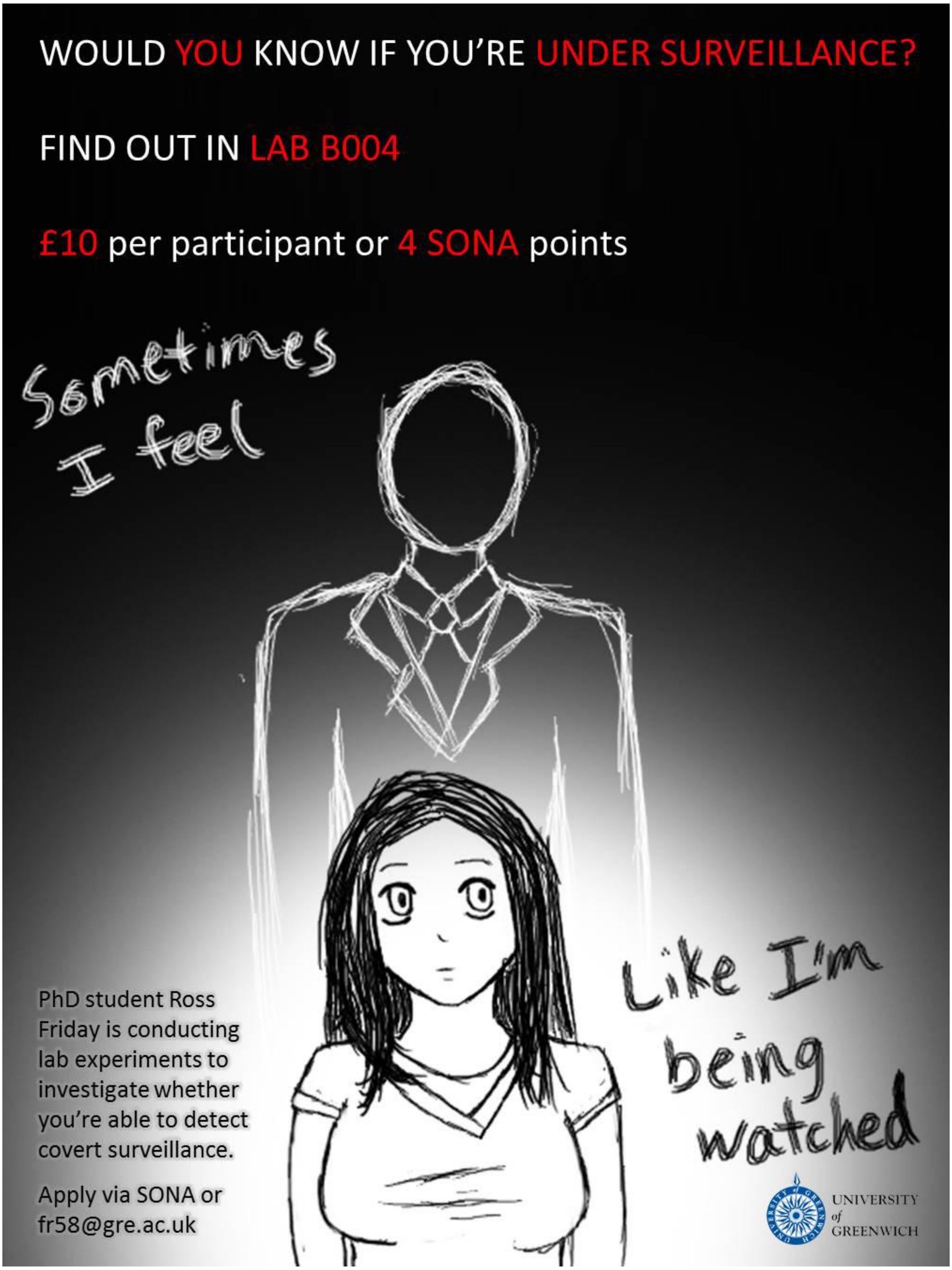
Yin, J., Yuan, J., Arfaei, N., Catalano, P., Allen, J., & Spengler, J. (2020). Effects of biophilic indoor environment on stress and anxiety recovery: A between-subjects experiment in virtual reality. *Environment International March*, 136.

Zdrenka, M., & Wilson, M.S. (2017). Individual difference correlates of psi performance in forced-choice precognition experiments: A meta-analysis (1945–2016), *Journal of Parapsychology*, *81*(1), 9-32.

7 APPENDICES

APPENDIX A.

Poster used for recruitment



APPENDIX B.

SPSS output, Study One, participant descriptive statistics

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age	112	18	68	31.74	12.992
Valid N (listwise)	112				

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	66	58.9	58.9	58.9
	Male	46	41.1	41.1	100.0
	Total	112	100.0	100.0	

APPENDIX C.

AEI Questionnaire

Anomalous Experiences Inventory (AEI) questionnaire.

Answer True / False

1. At times, my consciousness feels expanded beyond my body
2. I often have déjà vu experiences
3. I often seem to become aware of events before they happen
4. I often have psychic experiences
5. There have been events that I dreamed about before the event occurred
6. I have attended séances
7. I have had a near death experience
8. I often know what others are feeling or thinking without them telling me
9. I have experienced other planes of existence beyond the physical
10. I have experienced my body or objects floating in the air (levitation)
11. I have had a psychic experience

12. I use aouija board on a regular basis
13. I have experienced objects appearing or disappearing around me
14. I have had a mystical experience
15. I have had an out-of-body experience
16. I have had memories of a past life
17. I have communicated with the dead
18. I have seen a ghost or apparition
19. I have had the experience of time standing still
20. At times, I have felt possessed by an outside force
21. I have experienced or met an extra-terrestrial
22. I visit fortune tellers, palm readers, tarot card readers, or astrologers
23. My bizarre predictions have often come true
24. My horoscope is fairly accurate
25. I have had waking visions of an event which subsequently occurred
26. I have had a psychic or mystical experience
27. I have seen elves, fairies, and other types of little people
28. I have seen a UFO
29. I have experimented with witchcraft or sorcery
30. I believe that mind can control matter
31. I believe in life after death
32. I believe I have great power and energy within me waiting to be awakened
33. I want to understand the further reaches of my mind
34. I believe that many paranormal occurrences are real
35. I feel my mind can expand beyond its usual boundaries
36. I believe in the unconscious
37. I believe in reincarnation
38. I have lived before
39. I believe there is intelligent life on other planets
40. I believe that people have energy (an aura) surrounding their bodies

APPENDIX D.

CPES Questionnaire

Complex Partial Epileptic Signs (CPES) and Temporal Lobe Symptoms (TLS) subscales of the Personal Philosophy Inventory.

Directions: Read each of the following items, decide how you feel about it, and then mark your answer on the line next to the question.. If you agree with a statement, or feel that it is true about you, mark a "T" [W?] on the line. If you disagree with a statement, or feel that it is not true about you, mark an "F" on the line. Please try to answer every item.

1) Sometimes an event will occur that has special significance for me only.
2) There have been times when I have stared at an object and it appeared to become larger and larger.
3) While sitting quietly, I have had uplifting sensations as if I were driving quickly over a rolling road.
4) When I have a tough decision to make, a sign will be given and I will know what to do.
5) When relaxed or just before falling asleep, I sometimes feel pleasant vibrations moving through my whole body.
6) Sometimes in the early morning hours between midnight and 4:00 a.m., my experiences are very meaningful.
7) I have had a vision.
8) I have heard an inner voice call my name.
9) At least once a month, I experience intense smells that do not have an obvious source.
10) I use "hunches" more than simple learning to solve new problems.
11) I often feel as if things are not real.
12) When I walk upstairs, I sometimes note a strange smell from nowhere.
13) Once, in a crowded area, I suddenly could not recognize where I was.
14) I have had experiences when I felt as if I were somewhere else.
15) At least once a year I have dreams of floating or flying through the air.
16) At least once in my life, just before falling down, I have had the intense sensation of a smell from childhood.

APPENDIX E.

O-LIFE Questionnaire

The Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE).

Unusual Experiences (12 items)

1. When in the dark do you often see shapes and forms even though there is nothing there?
2. Are your thoughts sometimes so strong that you can almost hear them?
3. Have you ever thought that you had special, almost magical powers?
4. Have you sometimes sensed an evil presence around you, even though you could not see it?
5. Do you think that you could learn to read other's minds if you wanted to?
6. When you look in the mirror does your face sometimes seem quite different from usual?
7. Do ideas and insights sometimes come to you so fast that you cannot express them all?
8. Can some people make you aware of them just by thinking about you?
9. Does a passing thought ever seem so real it frightens you?
10. Do you feel that your accidents are caused by mysterious forces?
11. Do you ever have a sense of vague danger or sudden dread for reasons that you do not understand?
12. Does your sense of smell sometimes become unusually strong?

Cognitive Disorganisation (11 items)

13. Are you easily confused if too much happens at the same time?
14. Do you frequently have difficulty in starting to do things?
15. Are you a person whose mood goes up and down easily?
16. Do you dread going into a room by yourself where other people have already gathered and are talking?
17. Do you find it difficult to keep interested in the same thing for a long time?
18. Do you often have difficulties in controlling your thoughts?
19. Are you easily distracted from work by daydreams?
20. Do you ever feel that your speech is difficult to understand because the words are all mixed up and don't make sense?
21. Are you easily distracted when you read or talk to someone?
22. Is it hard for you to make decisions?
23. When in a crowded room, do you often have difficulty in following a conversation?

Introvertive Anhedonia (10 items)

24. Are there very few things that you have ever enjoyed doing?
25. Are you much too independent to get involved with other people?
26. Do you love having your back massaged?a
27. Do you find the bright lights of a city exciting to look at?a
28. Do you feel very close to your friends?a
29. Has dancing or the idea of it always seemed dull to you?
30. Do you like mixing with people?a
31. Is trying new foods something you have always enjoyed?a
32. Have you often felt uncomfortable when your friends touch you?
33. Do you prefer watching television to going out with people?

Impulsive Nonconformity (10 items)

34. Do you consider yourself to be pretty much an average sort of person?a
35. Would you like other people to be afraid of you?
36. Do you often feel the impulse to spend money which you know you can't afford?
37. Are you usually in an average kind of mood, not too high and not too low?a
38. Do you at times have an urge to do something harmful or shocking?
39. Do you stop to think things over before doing anything?a
40. Do you often overindulge in alcohol or food?
41. Do you ever have the urge to break or smash things?
42. Have you ever felt the urge to injure yourself?
43. Do you often feel like doing the opposite of what other people suggest even though you know they are right?

APPENDIX F.

Condition randomisation sequences

SEQUENCE 1: ORDER		**SEQUENCE 2: ORDER**		**SEQUENCE 3: ORDER**		**SEQUENCE 4: ORDER**	
Watched	30 secs	Both	30 secs	Listened to	30 secs	Both	30 secs
None	30 secs	Watched	30 secs	None	30 secs	Watched	30 secs
Both	30 secs	Listened to	30 secs	Listened to	30 secs	Listened to	30 secs
Watched	30 secs	Listened to	30 secs	Both	30 secs	None	30 secs
Listened to	30 secs	Watched	30 secs	Watched	30 secs	Listened to	30 secs
Both	30 secs	None	30 secs	Both	30 secs	None	30 secs
Listened to	30 secs	Both	30 secs	None	30 secs	Both	30 secs
None	30 secs	None	30 secs	Watched	30 secs	Watched	30 secs
Watched	30 secs	Both	30 secs	Listened to	30 secs	Both	30 secs
None	30 secs	Watched	30 secs	None	30 secs	Watched	30 secs
Both	30 secs	Listened to	30 secs	Listened to	30 secs	Listened to	30 secs
Watched	30 secs	Listened to	30 secs	Both	30 secs	None	30 secs
Listened to	30 secs	Watched	30 secs	Watched	30 secs	Listened to	30 secs
Both	30 secs	None	30 secs	Both	30 secs	None	30 secs
Listened to	30 secs	Both	30 secs	None	30 secs	Both	30 secs
None	30 secs	None	30 secs	Watched	30 secs	Watched	30 secs

APPENDIX G.

Consent Form

INFORMATION AND CONSENT

TITLE: Individual Differences in Psychosocial and Neurological Predictors of Surveillance Detection via Extrasensory Means

This study is part of my PhD research being conducted at the Department of Psychology and Counselling, The University of Greenwich, Avery Hill Road, Eltham, London, SE9 2UG.

PURPOSE OF STUDY:

Extrasensory detection of attention has been restricted to scopaesthesia (Sheldrake, 2003) - a phenomenon in which people respond via non-conventional means to being the subject of another persons' gaze. However this new research will incorporate the sense of being *listened to* as well as seen, which will be gauged by participants' reports of being watched and/or listened to, behavioural differences during surveillance, and their physiological reactions.

WHAT TO EXPECT:

You should only take part if you are over 18 years of age. The study will investigate which psychosocial and neurological factors may predict individuals' ability to 'sense' being watched and/or listened to. Before you start the experiment you will:

a) Be asked via a questionnaire your demographic information (age and gender).
b) Complete a version of the Anomalous Experiences Inventory (Gallagher et al., 1994) which will record your paranormal experiences and belief.
c) Undertake the O-LIFE questionnaire (Mason et al., 2005) to measure your levels of schizotypy.
d) Complete the CPES and TLS subscales of the Personal Philosophy Inventory (Persinger, 1983) as your temporal lobe is involved in hearing, speech and vision.

You will undergo three experiments as the 'detector' participant during which your ability to detect being watched and/or listened to will be measured by physiological fluctuations in electro dermal activity (EDA) which involves measuring your skin conductance, the accuracy of your self-reported 'feelings' of being under surveillance, and changes in your performance when taking a short computer based test to measure your responses to the Stroop effect (Ridley, 1935). These measurements will be taken during periods of:

a) No surveillance.
b) Being watched.
c) Being listened to.
d) Being watched and listened to.

Each participant will also undergo three experiments as the 'surveillance' participant during which they will observe and listen to another participant completing the experiments described above. The order will be decided randomly, and the entire experiment should last approximately an hour. Participants' ability to detect being watched and/or listened to will be analysed for possible relationships with the variables measured.

You are advised not to take part if you feel anxious about being tested. You may omit any questions that you do not wish to answer. The project is not expected to involve physical risks or mental discomfort or harm. Participants are free to withdraw themselves and their data from the study at any time they choose, with no consequence. Your participation is completely voluntary.

Forms and questionnaires will be kept securely and will be destroyed at the end of the study. All data kept in electronic format will be kept on a password accessed computer. The results of this research may be published or reported, but names will not be associated in any way with any published results.

To maintain confidentiality of records participants should invent a personal code, which will identify their data. This may be withdrawn anonymously by emailing the personal code to the address below. Data can be withdrawn at any time up until processing (30/12/2015).

Declaration: I have read and understood the above information. I have given been given a copy of this information. I have had an opportunity to ask questions and I have received satisfactory answers. I consent to participate in this study.

APPENDIX H.

Debrief Form

DEBRIEFING

Thank you for taking the time to participate in this study. The ability to detect attention has previously been restricted to the psychic staring effect, also known as scopaesthesia - a phenomenon in which people respond via non-conventional means to being the subject of another persons' gaze. However this new investigation is furthering the research by incorporating the sense of being heard as well as seen. The existence of these abilities will be gauged by the accuracy of participant's reports of being watched and/or listened to, differences in their behaviour during surveillance, and physiological reactions determined by electrodermal activity (EDA) which measures the electrical conductance of the participant's skin to indicate arousal.

Psychosocial and neurological factors which may be related to surveillance detection are also being investigated, and so participants' paranormal belief, temporal lobe lability and levels of schizotypy are being measured as both predictors of extrasensory ability, and as possible explanations for reported anomoulous experiences. The flexibility and permanence of your anomalous beliefs were also tested when you were shown falsified 'evidence' of your extrasensory ability on the print outs – these were in fact fabricated to oppose your previously stated beliefs about your own extrasensory ability to see if scientific data to the contrary would alter your previous held beliefs.

PLEASE DO NOT REVEAL THIS ASPECT OF THE EXPERIMENT TO ANYONE AS THIS WILL INVALIDATE THE RESEARCH.

This original and unique fusion of parapsychological, anomalistic, neuroscientific, and social factors represents an essential and progressive step in understanding possible extrasensory ability, as well as its psychosocial and neurological relationships. If you have any further questions please do not hesitate to contact me. If taking part in this study has affected how you feel about your own beliefs or your resilience to influence then the following websites are useful sources of local support and information; www.nhs.co.uk or www.mind.co.uk. You can also talk to your G.P. or to the university counselling service if you are a University of Greenwich student

- please use the university portal for details or call 020 8331 9444 if you need to speak with someone urgently.

Initial data analysis will be completed by the end of November 2015 so any withdrawal requests must be made before then. Remember that all information given in this study is kept securely and confidentially, and only the lead researcher, myself and their supervisor have access to the completed questionnaires. Questionnaires can only be identified by personal code words known only to participants, so not only will your information not be revealed, but it cannot be identified.

Once again, thank you for taking the time to take part in this research.

Name of investigator: Ross Friday. Contact details of investigator: fr58@gre.ac.uk

Contact details of supervisor: Dr David Luke (D.P.Luke@greenwich.ac.uk)

If you would like to be invited to take part in further research please insert your email address below. This will be kept separate from your responses.

APPENDIX I.

Study One, Self-Report Graphs

Quantity of Overall Surveillance Reports Observed in all Trials (Nt = 8) for all Participants (N = 112) Across Modes, Irrespective of Condition

Response mode	Sum of reported observations (all participants/ all conditions)	Mean	SD	% of trials where observation reported
Listened to (Nt=224)	463	4.13	1.57	51.67%
Watched (N=224)	487	4.35	1.41	54.35%
Overall	950	8.48	2.45	53.01%

The table above shows how often participants reported the feeling of being under surveillance across all conditions.

Quantity of 'Watched' Mode Responses Reported in Both Trials Combined (Nt =2) for all Participants (N = 112) for Each Condition

Condition	None	Listen	Watch	Both	Total
Mean	1.0268	1.1429	1.1250	1.0536	4.3482
Sum (MCE = 112)	115	128	126	118	487
Percentage (MCE = 50%)	51.34%	57.14%	56.25%	52.68%	54.35%

Quantity of 'Listened to' Responses Reported in Both Trials Combined (Nt =2) for all Participants (N = 112) for Each Condition

Condition	None	Listen	Watch	Both	Total
Mean	1.0446	.9196	.9643	1.2054	4.1339
Sum (MCE = 112)	117	103	108	135	463
Percentage (MCE = 50%)	52.23%	45.98%	48.21%	55.33%	51.67%

*Signal Detection Rates for the 'Listened to' Mode of Response for all Trials (Nt = 2) for all Participants (N = 112) for all of the Four Conditions **

	True	True %	False	False %
Positive	238 (hit)	53.1%	225(false alarm)	50.2%
Negative	223(correct rejection)	49.8%	210 (miss)	46.8%
Total		51.4%		48.6%

*[total number of trials across all conditions and all participants = 896]

Quantity of 'Watched and Listened to (both)' Responses Reported in Both Trials Combined (Nt =2) for all Participants (N = 112) for Each Condition

Condition	None	Listen	Watch	Both	Total
Mean	.6250	.6161	.5893	.6429	.6183
Sum (MCE = 56)	70	69	66	72	69.25
Percentage (MCE = 25%)	31.25%	30.80%	29.46%	32.14%	30.91%

Quantity of 'Neither Watched or Listened to (None)' Responses Reported in Both Trials Combined (Nt = 2) for all Participants (N = 112) for Each Condition

Condition	None	Listen	Watch	Both	Total
Mean	.5536	.5536	.5089	.3929	.5022
Sum (MCE = 56)	62	62	57	44	56.25
Percentage (MCE = 25%)	27.68%	27.68%	25.45%	19.64%	25.11%

APPENDIX J.

SPSS output, Study One, self-report Paired-Samples T-Tests

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Task_performance_while_under_NO_surveillance	1223.1608	112	268.15640	25.33840
	Task_performance_while_watched	1296.5188	112	289.97853	27.40040
Pair 2	Task_performance_while_under_NO_surveillance	1223.1608	112	268.15640	25.33840
	Task_performance_while_listened_to	1241.8387	112	228.14509	21.55768
Pair 3	Task_performance_while_under_NO_surveillance	1223.1608	112	268.15640	25.33840
	Task_performance_while_watched_AND_listened_to	1396.9412	112	427.63266	40.40749
Pair 4	Task_performance_while_watched	1296.5188	112	289.97853	27.40040
	Task_performance_while_listened_to	1241.8387	112	228.14509	21.55768

Pair 5	Task_performance_while_watched	1296.5188	112	289.97853	27.40040
	Task_performance_while_watched_AND_listened_to	1396.9412	112	427.63266	40.40749
Pair 6	Task_performance_while_listened_to	1241.8387	112	228.14509	21.55768
	Task_performance_while_watched_AND_listened_to	1396.9412	112	427.63266	40.40749

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Task_performance_while_under_NO_surveillance&Task_performance_while_watched	112	.164	.083
Pair 2	Task_performance_while_under_NO_surveillance&Task_performance_while_listened_to	112	.404	.000
Pair 3	Task_performance_while_under_NO_surveillance&Task_performance_while_watched_AND_listened_to	112	.169	.076
Pair 4	Task_performance_while_watched&Task_performance_while_listened_to	112	.372	.000
Pair 5	Task_performance_while_watched&Task_performance_while_watched_AND_listened_to	112	.256	.007
Pair 6	Task_performance_while_listened_to&Task_performance_while_watched_AND_listened_to	112	.403	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Task_performance_while_under_NO_surveillance - Task_performance_while_watched	-73.35804	361.16505	34.12689	-140.98275	-5.73333	-2.150	111	.034
Pair 2	Task_performance_while_under_NO_surveillance - Task_performance_while_listened_to	-18.67786	272.94238	25.79063	-69.78371	32.42800	-.724	111	.470
Pair 3	Task_performance_while_under_NO_surveillance - Task_performance_while_watched_AND_listened_to	-173.78036	464.90083	43.92900	-260.82860	-86.73211	-3.956	111	.000
Pair 4	Task_performance_while_watched - Task_performance_while_listened_to	54.68018	294.74339	27.85063	-.50771	109.86807	1.963	111	.052

Pair 5	Task_performance_while_watched - Task_performance_while_watched_AND_listened_to	-100.42232	451.17017	42.63157	-184.89963	-15.94501	-2.356	111	.020
Pair 6	Task_performance_while_listened_to - Task_performance_while_watched_AND_listened_to	-155.10250	395.42645	37.36429	-229.14233	-81.06267	-4.151	111	.000

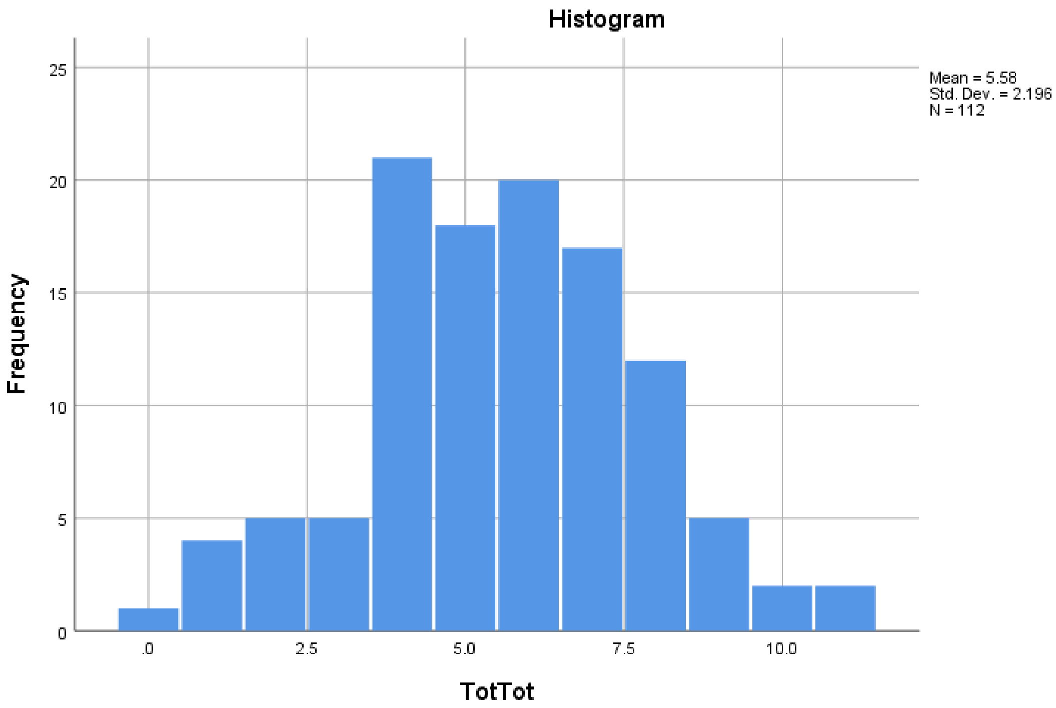
APPENDIX K.

SPSS output, Study One, self-reports tests of normality

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
TotTot	.102	112	.006	.977	112	.051

a. Lilliefors Significance Correction



APPENDIX L.

SPSS output, Study One, Individual Differences descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
AEI Experience	112	0	22	9.11	5.368
AEI Special Experience	112	0	24	10.42	5.568
AEI Belief	112	0	11	5.02	2.922
AEI Special Belief	112	0	13	6.43	3.368
Unusual Experiences	112	0	12	4.49	2.708
Cognitive Disorganisation	112	0	11	4.65	2.875
Introvertive Anhedonia	112	0	10	2.95	2.416
Impulsive Nonconformity	112	0	8	4.01	1.938
CPES	112	0	15	4.49	3.285
Valid N (listwise)	110				

Appendix M.

SPSS output, Study One, Pearson's correlation

analysis of individual differences with self-report ratio scores

		AEI Experience	AEI Special Experience	AEI Belief	AEI Special Belief
NoneSRRatio	Pearson Correlation	.112	.095	.151	.111
	Sig. (2-tailed)	.241	.320	.113	.243
	N	112	112	112	112
ListSRRatio	Pearson Correlation	-.090	-.103	-.130	-.141
	Sig. (2-tailed)	.346	.278	.172	.139
	N	112	112	112	112
WatchSRRatio	Pearson Correlation	-.047	-.038	-.076	-.062
	Sig. (2-tailed)	.623	.693	.429	.516
	N	112	112	112	112
BothSRRatio	Pearson Correlation	.094	.111	.064	.073
	Sig. (2-tailed)	.327	.244	.503	.442
	N	112	112	112	112

		Unusual Experiences	Cognitive Disorganisation	Introvertive Anhedonia	Impulsive Nonconformity	CPES
None SRRatio	Pearson Correlation	.209*	.125	.064	.046	.220*
	Sig. (2-tailed)	.027	.189	.502	.628	.020
	N	112	112	112	112	112
ListSR Ratio	Pearson Correlation	-.086	-.121	.280**	-.001	-.125
	Sig. (2-tailed)	.366	.205	.003	.988	.189
	N	112	112	112	112	112
WatchSR Ratio	Pearson Correlation	-.042	.034	-.106	-.057	-.007
	Sig. (2-tailed)	.663	.725	.264	.551	.944
	N	112	112	112	112	112
BothSR Ratio	Pearson Correlation	.245**	.131	.039	.070	.220*
	Sig. (2-tailed)	.009	.170	.681	.464	.019
	N	112	112	112	112	112

APPENDIX N.

SPSS output, Study One, Regression analysis of individual differences

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	T	Sig.
1	(Constant)	.301	.043		7.016	.000
	MeanIntrov	-.064	.054	-.120	-1.180	.241
	MeanCogDis	.048	.067	.097	.716	.475
	MeanAEIBelief	-.059	.210	-.122	-.280	.780
	MeanUnusual	-.095	.119	-.165	-.794	.429
	MeanImpulsive	-.047	.072	-.070	-.649	.518
	AEI_Belief_Score_inc_last_two_scores	-.001	.017	-.037	-.083	.934
	AEI_Experience_Score_inc_last_two_scores	.013	.017	.539	.736	.464
	AEImean	-.360	.492	-.515	-.731	.466
	MeanCPES	.121	.119	.193	1.017	.311

a. Dependent Variable: WatchSRRatio

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	.153	9	.017	1.405	.196[b]
	Residual	1.238	102	.012		
	Total	1.391	111			

a. Dependent Variable: ListSRRatio

b. Predictors: (Constant), MeanCPES, MeanIntrov, MeanImpulsive, MeanCogDis, AEI_Belief_Score_inc_last_two_scores, AEImean, MeanUnusual, MeanAEIBelief, AEI_Experience_Score_inc_last_two_scores

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	.239	.036		6.686	.000
	MeanIntrov	.128	.045	.276	2.826	.006
	MeanCogDis	-.046	.056	-.109	-.832	.407
	MeanAEIBelief	.046	.175	.111	.264	.792
	MeanUnusual	.123	.099	.248	1.241	.217
	MeanImpulsive	.009	.060	.015	.144	.886
	AEI_Belief_Score_inc_last_two_scores	-.005	.014	-.156	-.363	.718
	AEI_Experience_Score_inc_last_two_scores	-.004	.014	-.194	-.275	.784
	AEImean	.104	.410	.172	.254	.800
	MeanCPES	-.128	.099	-.236	-1.294	.199

a. Dependent Variable: ListSRRatio

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	.051	9	.006	1.221	.291[b]
	Residual	.476	102	.005		
	Total	.527	111			

a. Dependent Variable: BothSRRatio

b. Predictors: (Constant), MeanCPES, MeanIntrov, MeanImpulsive, MeanCogDis, AEI_Belief_Score_inc_last_two_scores, AEImean, MeanUnusual, MeanAEIBelief, AEI_Experience_Score_inc_last_two_scores

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	T	Sig.
1	(Constant)	.047	.022		2.129	.036
	MeanIntrov	.014	.028	.048	.492	.624
	MeanCogDis	-.008	.034	-.029	-.223	.824
	MeanAEIBelief	-.028	.109	-.110	-.260	.795
	MeanUnusual	.094	.062	.309	1.533	.128
	MeanImpulsive	-.014	.037	-.039	-.375	.709
	AEI_Belief_Score_inc_last_two_scores	-.002	.009	-.107	-.246	.806
	AEI_Experience_Score_inc_last_two_scores	.009	.009	.699	.982	.328
	AEImean	-.275	.254	-.740	-1.083	.281
	MeanCPES	.052	.061	.156	.846	.400

a. Dependent Variable: BothSRRatio

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.061	9	.007	1.190	.309[b]
	Residual	.577	102	.006		
	Total	.637	111			

a. Dependent Variable: NoneSRRatio

b. Predictors: (Constant), MeanCPES, MeanIntrov, MeanImpulsive, MeanCogDis, AEI_Belief_Score_inc_last_two_scores, AEImean, MeanUnusual, MeanAEIBelief, AEI_Experience_Score_inc_last_two_scores

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	T	Sig.
1	(Constant)	.053	.024		2.173	.032
	MeanIntrov	.024	.031	.077	.778	.438
	MeanCogDis	.010	.038	.035	.262	.794
	MeanAEIBelief	.143	.119	.507	1.194	.235
	MeanUnusual	.053	.068	.158	.782	.436
	MeanImpulsive	-.011	.041	-.029	-.280	.780
	AEI_Belief_Score_inc_last_two_scores	-.011	.010	-.497	-1.144	.255
	AEI_Experience_Score_inc_last_two_scores	-.010	.010	-.740	-1.039	.301
	AEImean	.269	.280	.657	.960	.339
	MeanCPES	.048	.068	.132	.716	.475

a. Dependent Variable: NoneSRRatio

APPENDIX O:

SPSS output, Study One, EDA descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
NoneEDAMean	112	.74	17.63	6.5350	2.98170
ListEDAMean	112	.76	18.04	6.5833	3.17126
WatchEDAMean	112	.72	16.23	6.4770	2.94860
BothEDAMean	112	.72	12.42	6.3892	2.85711
Valid N (listwise)	112				

APPENDIX P.

SPSS output, Study One, EDA ANOVA

Descriptive Statistics

	Mean	Std. Deviation	N
NoneEDAMean	6.5350	2.98170	112
WatchEDAMean	6.4770	2.94860	112
ListEDAMean	6.5833	3.17126	112
BothEDAMean	6.3892	2.85711	112

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Surveillance	Pillai's Trace	.028	1.051[b]	3.000	109.000	.373	.028
	Wilks' Lambda	.972	1.051[b]	3.000	109.000	.373	.028
	Hotelling's Trace	.029	1.051[b]	3.000	109.000	.373	.028
	Roy's Largest Root	.029	1.051[b]	3.000	109.000	.373	.028

a. Design: Intercept
Within Subjects Design: surveillance

b. Exact statistic

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	Df	Sig.	Epsilon[b]	
					Greenhouse-Geisser	Huynh-Feldt
surveillance	.759	30.289	5	.000	.831	.852

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
surveillance	Sphericity Assumed	2.342	3	.781	1.719	.163
	Greenhouse-Geisser	2.342	2.493	.939	1.719	.173
	Huynh-Feldt	2.342	2.555	.917	1.719	.171
	Lower-bound	2.342	1.000	2.342	1.719	.193
Error(surveillance)	Sphericity Assumed	151.237	333	.454		
	Greenhouse-Geisser	151.237	276.768	.546		
	Huynh-Feldt	151.237	283.633	.533		
	Lower-bound	151.237	111.000	1.362		

Tests of Within-Subjects Contrasts

Measure: NoneEDA

Source	Surveillance	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
surveillance	Linear	.613	1	.613	1.488	.225	.013
	Quadratic	.518	1	.518	1.419	.236	.013
	Cubic	1.210	1	1.210	2.069	.153	.018
Error(surveillance)	Linear	45.733	111	.412			
	Quadratic	40.565	111	.365			
	Cubic	64.939	111	.585			

Tests of Between-Subjects Effects

Measure: NoneEDA

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	18905.517	1	18905.517	548.905	.000	.832
Error	3823.087	111	34.442			

APPENDIX Q.

SPSS output, Study One, EDA_d descriptive statistics

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ListEDAratio	112	-.37	.48	.0039	.12705
WatchEDAratio	112	-.38	.32	.0010	.13512
BothEDAratio	112	-.47	.45	-.0117	.14241
Valid N (listwise)	112				

APPENDIX R.

SPSS output, Study One, EDA_d Pearson's correlation analysis of individual differences

		AEI Experience	AEI Special Experience	AEI Belief	AEI Special Belief
ListEDAratio	Pearson Correlation	-.045	-.033	.059	.058
	Sig. (2-tailed)	.638	.728	.535	.542
	N	112	112	112	112
WatchEDAratio	Pearson Correlation	-.030	-.039	-.008	-.006
	Sig. (2-tailed)	.752	.679	.931	.950
	N	112	112	112	112
BothEDAratio	Pearson Correlation	.044	.048	.003	.006
	Sig. (2-tailed)	.641	.616	.974	.946
	N	112	112	112	112

		Unusual Experiences	Cognitive Disorganisation	Introvertive Anhedonia	Impulsive Nonconformity	CPES
ListEDAratio	Pearson Correlation	.073	.042	.084	-.010	.063
	Sig. (2-tailed)	.442	.657	.378	.919	.506
	N	112	112	112	112	112
WatchEDAratio	Pearson Correlation	-.003	-.053	.106	.120	-.006
	Sig. (2-tailed)	.973	.582	.268	.208	.950
	N	112	112	112	112	112
BothEDAratio	Pearson Correlation	.064	.072	.121	.151	.097
	Sig. (2-tailed)	.502	.450	.203	.112	.307
	N	112	112	112	112	112

APPENDIX S.

SPSS output, Study One, Regression analysis of individual differences EDA_d

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	.064	9	.007	.419	.922[b]
	Residual	1.728	102	.017		
	Total	1.792	111			

a. Dependent Variable: ListEDAratio

b. Predictors: (Constant), AEI_Experience_Score_inc_last_two_scores, MeanIntrov, MeanCogDis, MeanImpulsive, MeanAEIBelief, MeanCPES, MeanUnusual, AEI_Belief_Score_inc_last_two_scores, AEImean

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	T	Sig.
1	(Constant)	-.011	.042		-.260	.796
	MeanAEIBelief	.014	.207	.029	.066	.948
	MeanUnusual	.084	.117	.149	.714	.477
	MeanCogDis	-.031	.066	-.065	-.479	.633
	MeanIntrov	.046	.053	.087	.855	.394
	MeanImpulsive	-.023	.071	-.036	-.328	.743
	MeanCPES	.035	.117	.058	.302	.763
	AEImean	-.353	.484	-.515	-.729	.468
	AEI_Belief_Score_inc_last_two_scores	.002	.017	.042	.093	.926
	AEI_Experience_Score_inc_last_two_scores	.007	.017	.326	.444	.658

a. Dependent Variable: ListEDAratio

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	.089	9	.010	.523	.855[b]
	Residual	1.937	102	.019		
	Total	2.026	111			

a. Dependent Variable: WatchEDAratio

b. Predictors: (Constant), AEI_Experience_Score_inc_last_two_scores, MeanIntrov, MeanCogDis, MeanImpulsive, MeanAEIBelief, MeanCPES, MeanUnusual, AEI_Belief_Score_inc_last_two_scores, AEImean

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-.021	.045		-.473	.638
	MeanAEIBelief	-.038	.219	-.076	-.174	.862
	MeanUnusual	.076	.124	.127	.612	.542
	MeanCogDis	-.084	.070	-.164	-1.204	.231
	MeanIntrov	.043	.057	.076	.752	.454
	MeanImpulsive	.114	.075	.165	1.524	.131
	MeanCPES	-.025	.124	-.039	-.204	.839
	AEImean	.259	.513	.356	.506	.614
	AEI_Belief_Score_inc_last_two_scores	.007	.018	.172	.384	.702
	AEI_Experience_Score_inc_last_two_scores	-.013	.018	-.517	-.706	.482

a. Dependent Variable: WatchEDAratio

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	.103	9	.011	.542	.840[b]
	Residual	2.148	102	.021		
	Total	2.251	111			

a. Dependent Variable: BothEDAratio

b. Predictors: (Constant), AEI_Experience_Score_inc_last_two_scores, MeanIntrov, MeanCogDis, MeanImpulsive, MeanAEIBelief, MeanCPES, MeanUnusual, AEI_Belief_Score_inc_last_two_scores, AEImean

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-.075	.047		-1.589	.115
	MeanAEIBelief	.054	.231	.102	.233	.816
	MeanUnusual	-.038	.131	-.060	-.289	.773
	MeanCogDis	.030	.073	.056	.412	.681
	MeanIntrov	.064	.060	.109	1.075	.285

	MeanImpulsive	.073	.079	.099	.920	.360
	MeanCPES	.084	.130	.122	.645	.520
	AEImean	-.348	.540	-.453	-.645	.521
	AEI_Belief_Score_inc_last_two_scores	-.009	.019	-.203	-.456	.649
	AEI_Experience_Score_inc_last_two_scores	.013	.019	.499	.683	.496

a. Dependent Variable: BothEDAratio

APPENDIX T.

SPSS output, Study One, task performance descriptive statistics

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
No Surveillance	112	864.58	2597.92	1223.1608	268.15640
Listened To	112	927.17	2239.75	1241.8387	228.14509
Watched	112	113.17	2687.50	1296.5188	289.97853
Watched and Listened to	112	941.50	3626.33	1396.9412	427.63266
Valid N (listwise)	112				

APPENDIX U.

SPSS output, Study One, task performance ANOVA

Descriptive Statistics

	Mean	Std. Deviation	N
Task_performance_while_watched	1296.5188	289.97853	112
Task_performance_while_listened_to	1241.8387	228.14509	112
Task_performance_while_watched_AND_listened_to	1396.9412	427.63266	112

Task_performance_while_under_NO_surveillance	1223.1608	268.15640	112

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
surveillance	Pillai's Trace	.156	6.697[b]	3.000	109.000	.000	.156
	Wilks' Lambda	.844	6.697[b]	3.000	109.000	.000	.156
	Hotelling's Trace	.184	6.697[b]	3.000	109.000	.000	.156
	Roy's Largest Root	.184	6.697[b]	3.000	109.000	.000	.156

a. Design: Intercept
Within Subjects Design: surveillance

b. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: Nonetaskperformance

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt
Surveillance	.625	51.645	5	.000	.784	.802

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
surveillance	Sphericity Assumed	2045714.338	3	681904.779	9.429	.000
	Greenhouse-Geisser	2045714.338	2.351	870275.750	9.429	.000
	Huynh-Feldt	2045714.338	2.405	850702.641	9.429	.000
	Lower-bound	2045714.338	1.000	2045714.338	9.429	.003

Error(surveillance)	Sphericity Assumed	24083136.678	333	72321.732		
	Greenhouse-Geisser	24083136.678	260.922	92300.056		
	Huynh-Feldt	24083136.678	266.926	90224.163		
	Lower-bound	24083136.678	111.000	216965.195		

Tests of Within-Subjects Contrasts

Measure: Nonetaskperformance

Source	Surveillance	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
surveillance	Linear	23639.335	1	23639.335	.377	.541	.003
	Quadratic	397175.871	1	397175.871	5.399	.022	.046
	Cubic	1624899.132	1	1624899.132	20.153	.000	.154
Error(surveillance)	Linear	6967806.848	111	62773.035			
	Quadratic	8165610.755	111	73564.061			
	Cubic	8949719.075	111	80628.100			

Tests of Between-Subjects Effects

Measure: Nonetaskperformance
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	745071713.251	1	745071713.251	4283.250	.000	.975
Error	19308458.985	111	173950.081			

APPENDIX V.

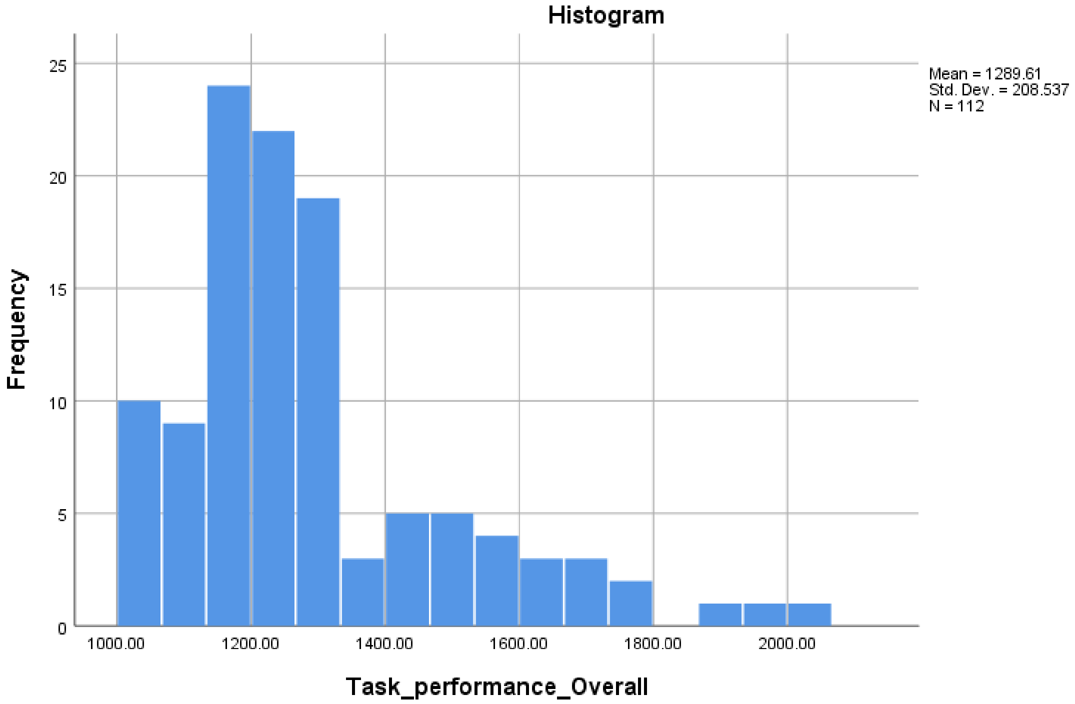
SPSS output, Study One, task performance tests normality

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	df	Sig.
Task_performance_Overall	.175	112	.000	.871	112	.000

a. Lilliefors Significance Correction

Task_Performance_Overall



APPENDIX W.

SPSS output, Study One, task performance paired t-tests

		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	None - Listened To	-18.68	272.94	25.79	-69.78	32.43	-.724	111	.470
Pair 2	None – Watched	-73.36	361.17	34.13	-140.98	-5.73	-2.150	111	.034
Pair 3	None - Both	-173.78	464.90	43.92	-260.82	-86.73	-3.956	111	.000

APPENDIX X.

SPSS output, Study One, task performance, Pearson's correlation with individual differences

		AEI Experience	AEI Special Experience	AEI Belief	AEI Special Belief
None	Pearson Correlation	-.241*	-.230*	.016	.003
	Sig. (2-tailed)	.010	.015	.871	.977
	N	112	112	112	112
Listened To	Pearson Correlation	-.092	-.095	.070	.067
	Sig. (2-tailed)	.332	.320	.461	.483
	N	112	112	112	112
Watched	Pearson Correlation	-.158	-.133	.028	.042
	Sig. (2-tailed)	.097	.163	.773	.663
	N	112	112	112	112
Both	Pearson Correlation	-.021	.006	.225*	.254**
	Sig. (2-tailed)	.823	.948	.017	.007
	N	112	112	112	112
	Sig. (2-tailed)	.000	.000	.000	
	N	112	112	112	112

		Unusual Experiences	Cognitive Disorganisation	Introvertive Anhedonia	Impulsive Nonconformity	CPES
No Surveillance	Pearson Correlation	.039	.143	.052	-.172	-.047
	Sig. (2-tailed)	.681	.132	.585	.069	.624
	N	112	112	112	112	112
Listened To	Pearson Correlation	.066	.182	.106	.051	.120
	Sig. (2-tailed)	.493	.054	.265	.590	.208
	N	112	112	112	112	112
Watched	Pearson Correlation	.026	.095	.107	-.115	.044
	Sig. (2-tailed)	.788	.317	.261	.227	.644
	N	112	112	112	112	112
Watched and Listened to	Pearson Correlation	.217*	.199*	.185	.071	.336**
	Sig. (2-tailed)	.022	.035	.051	.458	.000
	N	112	112	112	112	112
	Sig. (2-tailed)	.000	.000	.748	.000	
	N	112	112	112	112	112

APPENDIX Y.

SPSS output, Study One, task performance, Regression analysis

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	1035134.232	9	115014.915	1.414	.192[b]
	Residual	8298583.598	102	81358.663		
	Total	9333717.830	111			

a. Dependent Variable: Task_performance_while_watched
b. Predictors: (Constant), MeanCPES, MeanIntrov, MeanImpulsive, MeanCogDis, AEI_Belief_Score_inc_last_two_scores, AEImean, MeanUnusual, MeanAEIBelief, AEI_Experience_Score_inc_last_two_scores

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1306.969	92.626		14.110	.000
	AEI_Belief_Score_inc_last_two_scores	8.662	37.049	.101	.234	.816
	AEImean	-1652.260	1061.283	-1.056	-1.557	.123
	AEI_Experience_Score_inc_last_two_scores	40.311	36.761	.774	1.097	.275
	MeanAEIBelief	-21.584	453.130	-.020	-.048	.962
	MeanUnusual	12.694	257.080	.010	.049	.961
	MeanCogDis	49.772	144.058	.045	.346	.730
	MeanIntrov	142.306	117.059	.119	1.216	.227
	MeanImpulsive	-254.121	155.252	-.171	-1.637	.105
	MeanCPES	279.524	256.328	.199	1.090	.278

a. Dependent Variable: Task_performance_while_watched

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	546559.509	9	60728.834	1.184	.313[b]
	Residual	5231010.755	102	51284.419		
	Total	5777570.263	111			

a. Dependent Variable: Task_performance_while_listened_to

b. Predictors: (Constant), MeanCPES, MeanIntrov, MeanImpulsive, MeanCogDis, AEI_Belief_Score_inc_last_two_scores, AEImean, MeanUnusual, MeanAEIBelief, AEI_Experience_Score_inc_last_two_scores

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	1207.950	73.540		16.426	.000
	AEI_Belief_Score_inc_last_two_scores	7.845	29.415	.116	.267	.790
	AEImean	510.194	842.601	.414	.605	.546
	AEI_Experience_Score_inc_last_two_scores	-28.706	29.186	-.701	-.984	.328
	MeanAEIBelief	9.308	359.761	.011	.026	.979
	MeanUnusual	-117.907	204.107	-.117	-.578	.565
	MeanCogDis	114.493	114.374	.132	1.001	.319
	MeanIntrov	78.017	92.939	.083	.839	.403
	MeanImpulsive	19.576	123.261	.017	.159	.874
	MeanCPES	291.448	203.510	.264	1.432	.155

a. Dependent Variable: Task_performance_while_listened_to

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5771929.052	9	641325.450	4.503	.000[b]
	Residual	14526606.561	102	142417.711		
	Total	20298535.613	111			

a. Dependent Variable: Task_performance_while_watched_AND_listened_to

b. Predictors: (Constant), MeanCPES, MeanIntrov, MeanImpulsive, MeanCogDis, AEI_Belief_Score_inc_last_two_scores, AEImean, MeanUnusual, MeanAEIBelief, AEI_Experience_Score_inc_last_two_scores

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1176.358	122.550		9.599	.000
	AEI_Belief_Score_inc_last_two_scores	63.686	49.019	.502	1.299	.197
	AEImean	-2109.911	1404.143	-.914	-1.503	.136
	AEI_Experience_Score_inc_last_two_scores	36.454	48.636	.475	.750	.455
	MeanAEIBelief	-426.126	599.519	-.269	-.711	.479
	MeanUnusual	-219.809	340.132	-.116	-.646	.520
	MeanCogDis	-68.913	190.598	-.042	-.362	.718
	MeanIntrov	306.954	154.877	.174	1.982	.050
	MeanImpulsive	-96.662	205.407	-.044	-.471	.639
	MeanCPES	1246.329	339.138	.602	3.675	.000

. Dependent Variable: Task_performance_while_watched_AND_listened_to

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	1337721.346	9	148635.705	2.282	.022[b]
	Residual	6644050.611	102	65137.751		
	Total	7981771.957	111			

a. Dependent Variable: Task_performance_while_under_NO_surveillance

b. Predictors: (Constant), MeanCPES, MeanIntrov, MeanImpulsive, MeanCogDis, AEI_Belief_Score_inc_last_two_scores, AEImean, MeanUnusual, MeanAEIBelief, AEI_Experience_Score_inc_last_two_scores

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1306.173	82.879		15.760	.000
	AEI_Belief_Score_inc_last_two_scores	-16.416	33.151	-.206	-.495	.622

	AEImean	-661.827	949.611	-.457	-.697	.487
	AEI_Experience_Score_inc_last_two_scores	2.334	32.892	.048	.071	.944
	MeanAEIBelief	325.132	405.450	.327	.802	.424
	MeanUnusual	374.283	230.029	.315	1.627	.107
	MeanCogDis	84.916	128.900	.083	.659	.512
	MeanIntrov	85.470	104.742	.077	.816	.416
	MeanImpulsive	-275.619	138.915	-.200	-1.984	.050
	MeanCPES	-116.266	229.356	-.090	-.507	.613

a. Dependent Variable: Task_performance_while_under_NO_surveillance

APPENDIX Z.

SPSS output, Study Two, participant descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age	110	18	67	37.35	12.491
Valid N (listwise)	110				

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	53	48.2	48.2	48.2
	Female	57	51.8	51.8	100.0
	Total	110	100.0	100.0	

APPENDIX AA.

SPSS output, Study Two, EDA descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
EDAnone	109	.61	11.52	5.4374	2.80540
EDAlisten	109	.60	11.59	5.4357	2.78884

EDAwatch	110	.61	11.47	5.4204	2.85198
EDAboth	110	.62	11.92	5.4234	2.88026
Valid N (listwise)	109				

APPENDIX BB.

SPSS output, Study Two, EDA ANOVA

Descriptive Statistics

	Mean	Std. Deviation	N
EDAnone	5.4374	2.80540	109
EDAlisten	5.4357	2.78884	109
EDAwatch	5.4404	2.85739	109
EDAboth	5.4426	2.88648	109

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
EDA	Pillai's Trace	.000	.006[b]	3.000	106.000	.999	.000
	Wilks' Lambda	1.000	.006[b]	3.000	106.000	.999	.000
	Hotelling's Trace	.000	.006[b]	3.000	106.000	.999	.000
	Roy's Largest Root	.000	.006[b]	3.000	106.000	.999	.000

a. Design: Intercept
Within Subjects Design: EDA
b. Exact statistic

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Epsilon[b] Huynh-Feldt
EDA	.588	56.718	5	.000	.755	.772

Tests of Within-Subjects Effects

Measure: noneEDA

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
EDA	Sphericity Assumed	.003	3	.001	.006	.999	.000
	Greenhouse-Geisser	.003	2.265	.001	.006	.997	.000
	Huynh-Feldt	.003	2.316	.001	.006	.997	.000
	Lower-bound	.003	1.000	.003	.006	.939	.000
Error(EDA)	Sphericity Assumed	57.672	324	.178			
	Greenhouse-Geisser	57.672	244.583	.236			
	Huynh-Feldt	57.672	250.092	.231			
	Lower-bound	57.672	108.000	.534			

Tests of Within-Subjects Contrasts

Measure: noneEDA

Source	EDA	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
EDA	Linear	.002	1	.002	.010	.919	.000
	Quadratic	.000	1	.000	.004	.950	.000
	Cubic	.000	1	.000	.002	.964	.000
Error(EDA)	Linear	23.557	108	.218			
	Quadratic	11.419	108	.106			
	Cubic	22.696	108	.210			

Tests of Between-Subjects Effects

Measure: noneEDA

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	12898.240	1	12898.240	408.039	.000	.791
Error	3413.917	108	31.610			

APPENDIX CC.

SPSS output, Study Two, EDA_d score descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
EDARatiolisten	110	-1.85	189.00	1.7165	18.02905
EDARatioWatch	110	-41.54	3.01	-.3747	4.01576
EDARatioBoth	110	-41.45	2.55	-.3717	3.99190
Valid N (listwise)	110				

APPENDIX DD.

SPSS output, Study Two, EDA individual differences descriptive statistics

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
AEI Experience	110	.00	1.00	.3320	.30361
AEI Special Experience	110	.00	1.00	.3522	.29736
AEI Belief Mean	110	.00	1.00	.4721	.32760
AEI Special Belief Mean	110	.00	1.00	.5185	.30762
Unusual Experiences	110	0	12	4.83	2.851
Cognitive Disorganisation	109	0	11	4.39	2.603
Introvertive Anhedonia	110	0	9	4.94	1.473
Impulsive Nonconformity	109	0	9	4.66	2.038
CPES	110	0	13	5.02	3.543
Valid N (listwise)	109				

APPENDIX EE.

SPSS output, Study Two, EDA_d Pearson's correlation analysis of individual differences

		AEI Experience	AEI Special Experience	AEI Belief	AEI Special Belief
EDA Ratio listen	Pearson Correlation	-0.005	-0.008	0.025	0.009
	Sig. (2-tailed)	0.959	0.934	0.799	0.923
	N	110	110	110	110
EDA Ratio Watch	Pearson Correlation	0.002	0.004	-0.032	-0.019
	Sig. (2-tailed)	0.984	0.966	0.737	0.845
	N	110	110	110	110
EDA Ratio Both	Pearson Correlation	0.007	0.008	-0.029	-0.015
	Sig. (2-tailed)	0.946	0.931	0.767	0.878
	N	110	110	110	110

		Unusual Experiences	Cognitive Disorganisation	Introvertive Anhedonia	Impulsive Nonconformity	CPES
EDA Ratio listen	Pearson Correlation	-0.024	0.024	0.009	-0.074	0.029
	Sig. (2-tailed)	0.803	0.807	0.929	0.445	0.76
	N	110	110	110	110	110

EDA Ratio Watch	Pearson Correlation	0.036	-0.025	0.011	0.086	-0.023
	Sig. (2-tailed)	0.706	0.792	0.912	0.376	0.808
	N	110	110	110	110	110
EDA Ratio Both	Pearson Correlation	0.037	-0.022	0.004	0.079	-0.02
	Sig. (2-tailed)	0.698	0.816	0.964	0.416	0.835
	N	110	110	110	110	110

Appendix FF.

SPSS output, Study Two, EDA_d Regression analysis of individual differences

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	38.222	9	4.247	.245	.987[b]
	Residual	1719.527	99	17.369		
	Total	1757.749	108			

a. Dependent Variable: EDARatioWatch

b. Predictors: (Constant), AEI_Belief_Score_inc_last_two_scores, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Introvertive_Anhedonia_Total, O_LIFE_Unusual_Experiences_Total, CPES_Score, AEI_experience_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.535	1.817		-.845	.400
	O_LIFE_Cognitive_Disorganisation_Total	-.043	.165	-.028	-.261	.795
	O_LIFE_Introvertive_Anhedonia_Total	.080	.294	.029	.272	.786
	CPES_Score	-.104	.187	-.092	-.559	.577
	O_LIFE_Impulsive_Nonconformity_Total	.149	.206	.075	.724	.471
	AEI_belief_mean	-4.643	7.317	-.374	-.634	.527

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	5.726	8.173		.701	.485
	O_LIFE_Cognitive_Disorganisation_Total	.121	.743	.017	.162	.871
	AEI_Belief_Score_inc_last_two_scores	.254	.604	.249	.420	.676

a. Dependent Variable: EDARatioWatch

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	630.645	9	70.072	.199	.994[b]
	Residual	34797.448	99	351.489		
	Total	35428.093	108			

a. Dependent Variable: EDARatiolisten

b. Predictors: (Constant), AEI_Belief_Score_inc_last_two_scores, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Introvertive_Anhedonia_Total, O_LIFE_Unusual_Experiences_Total, CPES_Score, AEI_experience_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
	O_LIFE_Impulsive_Nonconformity_Total	-.541	.925	-.061	-.585	.560
	AEI_belief_mean	21.996	32.917	.395	.668	.506
	O_LIFE_Unusual_Experiences_Total	-.597	1.018	-.094	-.586	.559
	AEI_experience_mean	7.026	82.889	.116	.085	.933
	AEI_Experience_Score_inc_last_two_scores_mean	-12.851	87.032	-.207	-.148	.883
	AEI_Belief_Score_inc_last_two_scores	-1.352	2.718	-.296	-.497	.620

a. Dependent Variable: EDARatiolisten

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	34.479	9	3.831	.223	.991[b]
	Residual	1702.104	99	17.193		
	Total	1736.583	108			

a. Dependent Variable: EDARatioBoth

b. Predictors: (Constant), AEI_Belief_Score_inc_last_two_scores, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Introvertive_Anhedonia_Total, O_LIFE_Unusual_Experiences_Total, CPES_Score, AEI_experience_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

1	(Constant)	-1.419	1.808		-.785	.434
	O_LIFE_Cognitive_Disorganisation_Total	-.037	.164	-.024	-.222	.825
	O_LIFE_Introvertive_Anhedonia_Total	.057	.292	.021	.195	.846
	CPES_Score	-.096	.186	-.085	-.517	.606
	O_LIFE_Impulsive_Nonconformity_Total	.133	.205	.068	.652	.516
	AEI_belief_mean	-4.877	7.280	-.395	-.670	.504
	O_LIFE_Unusual_Experiences_Total	.158	.225	.112	.700	.486
	AEI_experience_mean	.605	18.332	.045	.033	.974
	AEI_Experience_Score_inc_last_two_scores_mean	.688	19.249	.050	.036	.972
	AEI_Belief_Score_inc_last_two_scores	.280	.601	.276	.465	.643

a. Dependent Variable: EDARatioBoth

APPENDIX GG.

SPSS output, Study Two, EDA with Stroop descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
EDA_None_stroop	108	.62	7.62	3.6598	1.75263
EDA_Listen_stroop	108	.67	7.64	3.7840	1.59856
EDA_Watch_stroop	108	.49	7.32	4.0406	1.55006
EDA_Both_stroop	108	.50	7.47	3.9301	1.56506
Valid N (listwise)	108				

APPENDIX HH.

SPSS output, Study Two, EDA with Stroop ANOVA

Descriptive Statistics

	Mean	Std. Deviation	N
EDA_None_stroop	3.6598	1.75263	108
EDA_Listen_stroop	3.7840	1.59856	108
EDA_Watch_stroop	4.0406	1.55006	108
EDA_Both_stroop	3.9301	1.56506	108

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
EDArawscores	Pillai's Trace	.126	5.061[b]	3.000	105.000	.003	.126
	Wilks' Lambda	.874	5.061[b]	3.000	105.000	.003	.126
	Hotelling's Trace	.145	5.061[b]	3.000	105.000	.003	.126
	Roy's Largest Root	.145	5.061[b]	3.000	105.000	.003	.126

a. Design: Intercept
Within Subjects Design: EDArawscores
b. Exact statistic

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Epsilon[b] Huynh-Feldt
EDArawscores	.583	57.079	5	.000	.743	.759

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
EDArawscores	Sphericity Assumed	8.990	3	2.997	8.101	.000
	Greenhouse-Geisser	8.990	2.228	4.034	8.101	.000
	Huynh-Feldt	8.990	2.278	3.946	8.101	.000
	Lower-bound	8.990	1.000	8.990	8.101	.005

Error(EDArawscores)	Sphericity Assumed	118.734	321	.370		
	Greenhouse-Geisser	118.734	238.446	.498		
	Huynh-Feldt	118.734	243.751	.487		
	Lower-bound	118.734	107.000	1.110		

Tests of Within-Subjects Contrasts

Measure: NoneEDAStroop

Source	EDArawscores	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
EDArawscores	Linear	6.154	1	6.154	10.340	.002	.088
	Quadratic	1.488	1	1.488	6.630	.011	.058
	Cubic	1.349	1	1.349	4.647	.033	.042
Error(EDArawscores)	Linear	63.676	107	.595			
	Quadratic	24.007	107	.224			
	Cubic	31.052	107	.290			

Tests of Between-Subjects Effects

Measure: NoneEDAStroop
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	6415.415	1	6415.415	684.712	.000	.865
Error	1002.538	107	9.370			

APPENDIX II.

SPSS output, Study Two, EDA with Stroop paired samples t-tests

		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper	t	df	Sig. (2-tailed)
Pair 1	EDA_None_stroop EDA_Listen_stroop	-.12417	.82531	.07942	-.28160	.03326	-1.564	107	.121
Pair 2	EDA_None_stroop EDA_Watch_stroop	-.38083	1.05235	.10126	-.58157	-.18009	-3.761	107	.000
Pair 3	EDA_None_stroop EDA_Both_stroop	-.27028	1.05270	.10130	-.47109	-.06947	-2.668	107	.009

APPENDIX JJ.

SPSS output, Study Two, EDA_d with Stroop descriptive statistics

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
EDA_Listen_Subtracted_Differences_stroop	108	-2.75	3.71	.1242	.82531
EDA_Both_Subtracted_Differences_stroop	108	-3.99	4.22	.2703	1.05270
EDA_Watch_Subtracted_Differences_stroop	108	-3.21	4.44	.3808	1.05235
Valid N (listwise)	108				

APPENDIX KK.

SPSS output, Study Two, EDA_d with Stroop Pearson's Correlation analysis

		AEI_experience_mean	AEI_Experience_Score_inc_last_two_scores_mean	AEI_belief_mean	AEI_Belief_Score_inc_last_two_scores
EDARatiolisten	Pearson Correlation	-0.112	-0.108	-0.074	-0.048
	Sig. (2-tailed)	0.25	0.268	0.444	0.623
	N	108	108	108	108
EDARatioWatch	Pearson Correlation	-0.143	-0.132	-0.113	-0.106
	Sig. (2-tailed)	0.139	0.172	0.246	0.276
	N	108	108	108	108
EDARatioBoth	Pearson Correlation	-0.185	-0.178	-0.146	-0.13

	Sig. (2-tailed)	0.055	0.065	0.132	0.18
	N	108	108	108	108

		O_LIFE_Unusual_Experiences_Total	O_LIFE_Cognitive_Disorganisation_Total	O_LIFE_Introvertive_Anhedonia_Total	O_LIFE_Impulsive_Nonconformity_Total	CPES_Score
EDA Ratiolisten	Pearson Correlation	-0.077	-0.133	-0.042	0.105	-0.119
	Sig. (2-tailed)	0.429	0.169	0.666	0.278	0.22
	N	108	108	108	108	108
EDA Ratio Watch	Pearson Correlation	-0.117	-0.12	-0.041	0.123	-0.116
	Sig. (2-tailed)	0.226	0.215	0.676	0.205	0.231
	N	108	108	108	108	108
EDA Ratio Both	Pearson Correlation	-0.146	-0.068	-0.014	0.15	-0.143
	Sig. (2-tailed)	0.133	0.484	0.882	0.122	0.138
	N	108	108	108	108	108

APPENDIX LL.

SPSS output, Study Two, EDA_d with Stroop regression analysis

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	3.912	9	.435	.618	.779[b]
	Residual	68.969	98	.704		
	Total	72.881	107			

a. Dependent Variable: EDA_Listen_Subtracted_Differences_stroop

b. Predictors: (Constant), AEI_Belief_Score_inc_last_two_scores, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Introvertive_Anhedonia_Total, O_LIFE_Unusual_Experiences_Total, CPES_Score, AEI_experience_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.111	.366		.302	.763
	O_LIFE_Cognitive_Disorganisation_Total	-.038	.034	-.120	-1.133	.260
	O_LIFE_Introvertive_Anhedonia_Total	-.004	.059	-.007	-.067	.947
	O_LIFE_Impulsive_Nonconformity_Total	.040	.042	.100	.970	.334
	CPES_Score	-.014	.038	-.059	-.358	.721
	O_LIFE_Unusual_Experiences_Total	.002	.046	.007	.044	.965
	AEI_experience_mean	1.431	3.764	.517	.380	.705
	AEI_belief_mean	-1.355	1.486	-.531	-.912	.364
	AEI_Experience_Score_inc_last_two_scores_mean	-1.783	3.950	-.632	-.451	.653
	AEI_Belief_Score_inc_last_two_scores	.128	.122	.614	1.051	.296

a. Dependent Variable: EDA_Listen_Subtracted_Differences_stroop

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	7.408	9	.823	.726	.684[b]
	Residual	111.089	98	1.134		
	Total	118.496	107			

a. Dependent Variable: EDA_Watch_Subtracted_Differences_stroop

b. Predictors: (Constant), AEI_Belief_Score_inc_last_two_scores, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Introvertive_Anhedonia_Total, O_LIFE_Unusual_Experiences_Total, CPES_Score, AEI_experience_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.512	.464		1.103	.273
	O_LIFE_Cognitive_Disorganisation_Total	-.041	.043	-.101	-.958	.340
	O_LIFE_Introvertive_Anhedonia_Total	-.012	.075	-.017	-.163	.871
	O_LIFE_Impulsive_Nonconformity_Total	.063	.053	.123	1.194	.235
	CPES_Score	.004	.048	.014	.083	.934
	O_LIFE_Unusual_Experiences_Total	-.025	.058	-.069	-.434	.665
	AEI_experience_mean	-5.673	4.777	-1.607	-1.188	.238
	AEI_belief_mean	.903	1.885	.277	.479	.633
	AEI_Experience_Score_inc_last_two_scores_mean	5.641	5.013	1.568	1.125	.263
	AEI_Belief_Score_inc_last_two_scores	-.091	.155	-.343	-.590	.557

a. Dependent Variable: EDA_Watch_Subtracted_Differences_stroop

Coefficients[a]

Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	7.516	9	.835	.737	.674[b]
	Residual	111.060	98	1.133		
	Total	118.575	107			

a. Dependent Variable: EDA_Both_Subtracted_Differences_stroop

b. Predictors: (Constant), AEI_Belief_Score_inc_last_two_scores, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Introvertive_Anhedonia_Total, O_LIFE_Unusual_Experiences_Total, CPES_Score, AEI_experience_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

1	(Constant)	.207	.464		.446	.657
	O_LIFE_Cognitive_Disorganisation_Total	-.022	.043	-.053	-.505	.614
	O_LIFE_Introvertive_Anhedonia_Total	.007	.075	.011	.100	.921
	O_LIFE_Impulsive_Nonconformity_Total	.072	.053	.141	1.373	.173
	CPES_Score	.000	.048	.001	.007	.994
	O_LIFE_Unusual_Experiences_Total	-.030	.058	-.082	-.517	.606
	AEI_experience_mean	-2.256	4.776	-.639	-.472	.638
	AEI_belief_mean	-.312	1.885	-.096	-.165	.869
	AEI_Experience_Score_inc_last_two_scores_mean	1.910	5.013	.531	.381	.704
	AEI_Belief_Score_inc_last_two_scores	.022	.155	.083	.142	.887

a. Dependent Variable: EDA_Both_Subtracted_Differences_stroop

APPENDIX MM.

SPSS output, Study Two, task performance descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Stroop_Test_None	107	10529	24772	14510.17	2284.337
Stroop_Test_Listen	109	10400	26174	14548.04	2758.212
Stroop_Test_Watch	108	10700	25042	15020.78	3046.873
Stroop_Test_Both	108	9504	27893	15180.56	3226.619
Valid N (listwise)	102				

APPENDIX NN.

SPSS output, Study Two, task performance ANOVA

Descriptive Statistics

	Mean	Std. Deviation	N
Stroop_Test_None	14487.98	2257.719	102
Stroop_Test_Listen	14428.51	2713.988	102
Stroop_Test_Watch	15103.46	3059.360	102
Stroop_Test_Both	15183.38	3289.972	102

Effect		Value	F	Hypothesis df	Error df	Sig.
taskperformanceraw scores	Pillai's Trace	.062	2.169[b]	3.000	99.000	.096
	Wilks' Lambda	.938	2.169[b]	3.000	99.000	.096
	Hotelling's Trace	.066	2.169[b]	3.000	99.000	.096
	Roy's Largest Root	.066	2.169[b]	3.000	99.000	.096

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	Df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt
taskperformanceraw scores	.858	15.252	5	.009	.912	.940

Source		Type III Sum of Squares	Df	Mean Square	F	Sig.
taskperformanceraw scores	Sphericity Assumed	48391747.784	3	16130582.595	3.019	.030
	Greenhouse-Geisser	48391747.784	2.737	17677748.781	3.019	.035

	Huynh-Feldt	48391747.784	2.821	17152871.959	3.019	.033
	Lower-bound	48391747.784	1.000	48391747.784	3.019	.085
Error(taskperformancerawscores)	Sphericity Assumed	1618703520.216	303	5342255.842		
	Greenhouse-Geisser	1618703520.216	276.481	5854658.761		
	Huynh-Feldt	1618703520.216	284.942	5680825.841		
	Lower-bound	1618703520.216	101.000	16026767.527		

Source	taskperformancerawscores	Type III Sum of Squares	df	Mean Square	F	Sig.
Taskperformancerawscores	Linear	38882334.825	1	38882334.825	5.597	.020
	Quadratic	495469.422	1	495469.422	.097	.756
	Cubic	9013943.537	1	9013943.537	2.260	.136
Error(taskperformancerawscores)	Linear	701584365.975	101	6946379.861		
	Quadratic	514239287.578	101	5091478.095		
	Cubic	402879866.663	101	3988909.571		

Tests of Between-Subjects Effects

Measure: strooptestnone

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared

Intercept	89378384283.333	1	89378384283.333	5377.807	.000	.982
Error	1678605580.667	101	16619857.234			

APPENDIX OO.

SPSS output, Study Two, task performance paired t-tests

		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed
Pair 1	Stroop_Test_None Stroop_Test_Listen	-13.425	2831.348	275.005	-558.708	531.859	-.049	105	.96
Pair 2	Stroop_Test_None Stroop_Test_Watch	-508.600	3532.850	344.771	-1192.294	175.094	-1.48	104	.14
Pair 3	Stroop_Test_NoneS troop_Test_Both	-702.571	3430.281	334.761	-1366.415	-38.728	-2.10	104	.03

APPENDIX PP.

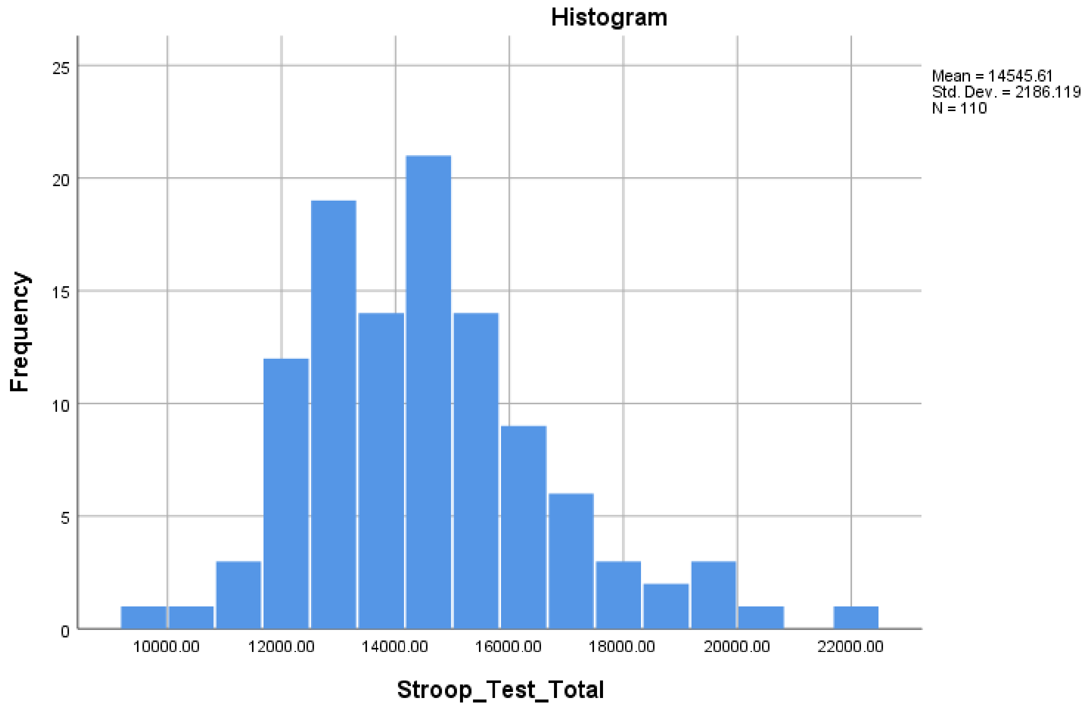
SPSS output, Study Two, task performance, tests of normality

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	df	Sig.
Stroop_Test_Total	.086	110	.045	.960	110	.002

a. Lilliefors Significance Correction

Stroop_Test_Total



APPENDIX QQ.

SPSS output, Study Two, task performance

Pearson's correlation analysis with individual differences

		AEI Experience	AEI Special Experience	AEI Belief	AEI Special Belief
Stroop_Test_None	Pearson Correlation	-.003	-.005	.043	.021
	Sig. (2-tailed)	.977	.960	.661	.827
	N	107	107	107	107

Stroop_Test_Listen	Pearson Correlation	.090	.103	.082	.084
	Sig. (2-tailed)	.351	.287	.397	.384
	N	109	109	109	109
Stroop_Test_Watch	Pearson Correlation	.140	.143	.072	.050
	Sig. (2-tailed)	.147	.139	.457	.609
	N	108	108	108	108
Stroop_Test_Both	Pearson Correlation	.058	.069	.065	.063
	Sig. (2-tailed)	.554	.477	.502	.515
	N	108	108	108	108

		Unusual Experiences	Cognitive Disorganisation	Introvertive Anhedonia	Impulsive Nonconformity	CPES
Stroop_Test None	Pearson Correlation	-.046	-.054	.053	-.119	-.041
	Sig. (2-tailed)	.636	.583	.587	.223	.674
	N	107	106	107	106	107
Stroop_Test Listen	Pearson Correlation	.095	-.117	.082	.092	.044
	Sig. (2-tailed)	.327	.228	.397	.346	.650
	N	109	108	109	108	109
Stroop_Test Watch	Pearson Correlation	.044	-.016	.054	-.132	.160
	Sig. (2-tailed)	.648	.867	.582	.176	.097
	N	108	107	108	107	108
Stroop_Test Both	Pearson Correlation	-.005	.036	.042	.036	.109
	Sig. (2-tailed)	.960	.714	.668	.710	.264
	N	108	107	108	107	108

APPENDIX RR.

SPSS output, Study Two, task performance Regression analysis

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	23457950.715	9	2606438.968	.473	.890[b]
	Residual	529294250.841	96	5513481.780		
	Total	552752201.557	105			

a. Dependent Variable: Stroop_Test_None
b. Predictors: (Constant), AEI_Belief_Score_inc_last_two_scores, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Introvertive_Anhedonia_Total, O_LIFE_Unusual_Experiences_Total, CPES_Score, AEI_experience_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	15046.737	1038.743		14.486	.000
	O_LIFE_Cognitive_Disorganisation_Total	-43.397	93.536	-.050	-.464	.644
	O_LIFE_Introvertive_Anhedonia_Total	112.616	167.741	.072	.671	.504
	CPES_Score	-37.233	105.545	-.058	-.353	.725
	O_LIFE_Impulsive_Nonconformity_Total	-102.711	117.681	-.091	-.873	.385
	AEI_belief_mean	5238.789	4219.898	.744	1.241	.217
	O_LIFE_Unusual_Experiences_Total	-46.215	128.523	-.058	-.360	.720
	AEI_experience_mean	-1164.508	10555.882	-.152	-.110	.912
	AEI_Experience_Score_inc_last_two_scores_mean	732.177	11084.559	.093	.066	.947
	AEI_Belief_Score_inc_last_two_scores	-352.435	348.132	-.613	-1.012	.314

a. Dependent Variable: Stroop_Test_None

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	61503794.943	9	6833754.994	.881	.545[b]
	Residual	760130804.270	98	7756436.778		
	Total	821634599.213	107			

a. Dependent Variable: Stroop_Test_Listen
b. Predictors: (Constant), AEI_Belief_Score_inc_last_two_scores, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Introvertive_Anhedonia_Total, O_LIFE_Unusual_Experiences_Total, CPES_Score, AEI_experience_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	13264.708	1249.319		10.618	.000
	O_LIFE_Cognitive_Disorganisation_Total	-139.972	112.513	-.130	-1.244	.216
	O_LIFE_Introvertive_Anhedonia_Total	191.202	196.557	.102	.973	.333
	CPES_Score	-63.508	125.901	-.082	-.504	.615
	O_LIFE_Impulsive_Nonconformity_Total	135.582	137.581	.099	.985	.327
	AEI_belief_mean	3842.247	4957.625	.453	.775	.440
	O_LIFE_Unusual_Experiences_Total	77.819	152.750	.080	.509	.612
	AEI_experience_mean	-17788.863	12371.704	-1.920	-1.438	.154
	AEI_Experience_Score_inc_last_two_scores_mean	20204.568	13006.217	2.136	1.553	.124
	AEI_Belief_Score_inc_last_two_scores	-409.694	407.426	-.588	-1.006	.317

a. Dependent Variable: Stroop_Test_Listen

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	90880097.621	9	10097788.625	1.086	.380[b]
	Residual	902256957.968	97	9301618.123		
	Total	993137055.589	106			

a. Dependent Variable: Stroop_Test_Watch
b. Predictors: (Constant), AEI_Belief_Score_inc_last_two_scores, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Introvertive_Anhedonia_Total, O_LIFE_Unusual_Experiences_Total, CPES_Score, AEI_experience_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	15520.453	1341.730		11.567	.000
	O_LIFE_Cognitive_Disorganisation_Total	-68.181	122.692	-.058	-.556	.580
	O_LIFE_Introvertive_Anhedonia_Total	103.177	218.557	.050	.472	.638
	CPES_Score	216.629	141.646	.253	1.529	.129
	O_LIFE_Impulsive_Nonconformity_Total	-149.248	150.563	-.100	-.991	.324
	AEI_belief_mean	4700.080	5385.981	.497	.873	.385
	O_LIFE_Unusual_Experiences_Total	-200.449	170.564	-.186	-1.175	.243
	AEI_experience_mean	-14962.678	13528.792	-1.445	-1.106	.271
	AEI_Experience_Score_inc_last_two_scores_mean	17981.690	14184.258	1.706	1.268	.208
	AEI_Belief_Score_inc_last_two_scores	-552.850	444.049	-.714	-1.245	.216

a. Dependent Variable: Stroop_Test_Watch

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	62170448.926	9	6907827.658	.637	.763[b]
	Residual	1051743243.317	97	10842713.849		
	Total	1113913692.243	106			

a. Dependent Variable: Stroop_Test_Both

b. Predictors: (Constant), AEI_Belief_Score_inc_last_two_scores, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Introvertive_Anhedonia_Total, O_LIFE_Unusual_Experiences_Total, CPES_Score, AEI_experience_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	14432.003	1438.665		10.032	.000
	O_LIFE_Cognitive_Disorganisation_Total	16.713	130.706	.014	.128	.899
	O_LIFE_Introvertive_Anhedonia_Total	28.079	233.158	.013	.120	.904
	CPES_Score	205.365	147.642	.227	1.391	.167
	O_LIFE_Impulsive_Nonconformity_Total	58.754	162.463	.037	.362	.718
	AEI_belief_mean	3329.003	5816.299	.330	.572	.568
	O_LIFE_Unusual_Experiences_Total	-249.017	178.957	-.221	-1.391	.167
	AEI_experience_mean	-21576.786	14652.226	-1.991	-1.473	.144
	AEI_Experience_Score_inc_last_two_scores_mean	23054.576	15356.442	2.081	1.501	.137

	AEI_Belief_Score_inc_last_two_scores	-312.864	480.421	-.377	-.651	.516

a. Dependent Variable: Stroop_Test_Both

APPENDIX SS.

Study Two, Self-Report Bias Graphs

Quantity of Overall Surveillance Reports Observed in Trials (Nt = 16) for all Participants (N = 112) Across Modes, Irrespective of Condition

Response mode	Sum of reported observations (all participants/ all conditions)	Mean	SD	% of trials where observation reported
Listen to (N=440)	848	7.71	3.45	48.18%
Watched (N=440)	922	8.38	3.22	52.39%
Overall	1770	8.05	3.33	50.29%

Note. Table32 above shows how often participants reported the feeling of surveillance across all conditions.

Quantity of 'Watched' Mode Responses Reported in Both Trials Combined (Nt = 4) for all Participants (N = 110) for Each Condition

Condition	None	Listen	Watch	Both	Total
Mean	1.84	2.06	2.26	2.26	8.42
Sum (MCE = 220)	202	227	249	249	927

Percentage (MCE = 50%)	45.91%	51.59%	56.59%	56.59%	52.67%

Quantity of 'Listened to' Responses Reported in Both Trials Combined (Nt =4) for all Participants (N = 110) for Each Condition

Condition	None	Listen	Watch	Both	Total
Mean	1.78	1.89	1.68	2.35	7.70
Sum (MCE = 220)	196	208	185	259	848
Percentage (MCE = 50%)	44.55%	47.27%	42.05%	58.86%	48.18%

Quantity of 'Watched and Listened to (Both)' Responses Reported in Both Trials Combined (Nt = 4) for all Participants (N = 112) for Each Condition

Condition	None	Listen	Watch	Both	Total
Mean	1.03	1.09	.99	1.45	1.14
Sum (MCE = 110)	113	120	109	160	125.50
Percentage (MCE = 25%)	25.68%	27.27%	24.77%	36.36%	28.52%

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
ResponseBias	109	16.0826	5.73993	.54979

One-Sample Test

	Test Value = 16					
					95% Confidence Interval of the Difference	
	T	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
ResponseBias	.150	108	.881	.08257	-1.0072	1.1723

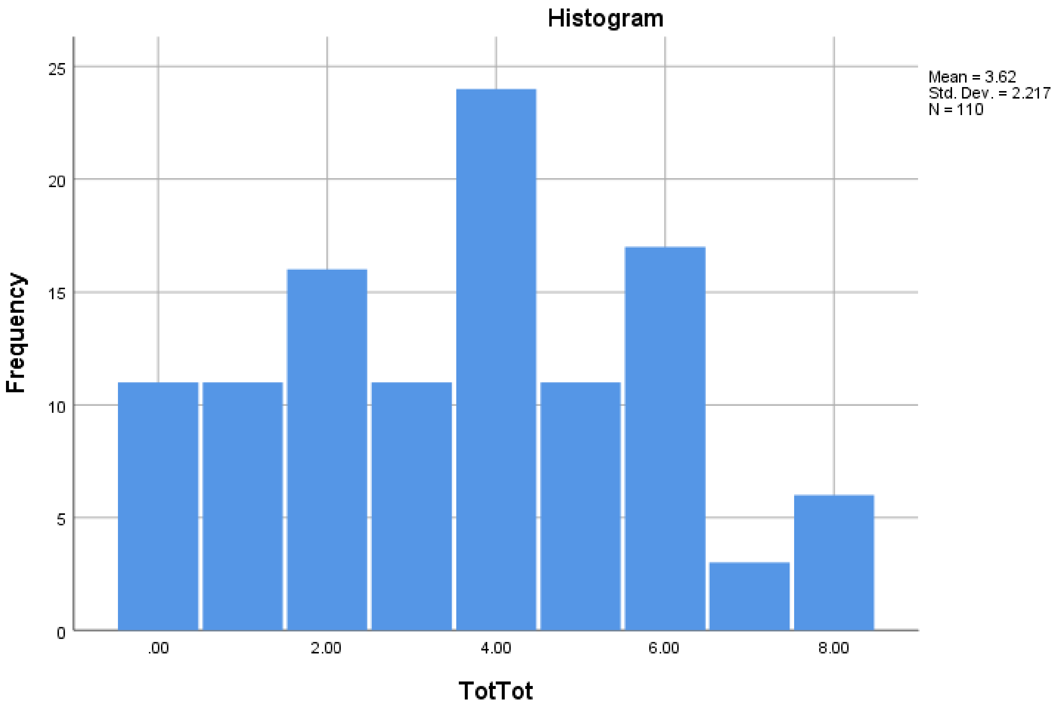
APPENDIX TT.

SPSS output, Study Two, self-reports tests of normality

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	Df	Sig.
TotTot	.123	110	.000	.952	110	.001

a. Lilliefors Significance Correction



APPENDIX UU.

Study Two Self-Report One-Sample T-Tests

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
TOTALtrue	5280	.4839	.49979	.00688

One-Sample Test

	Test Value = .4166666					
					95% Confidence Interval of the Difference	
	t	Df	Sig. (2-tailed)	Mean Difference	Lower	Upper
TOTALtrue	9.775	5279	.0000000000000000000002	.06723	.0538	.0807

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
TotalFalse	5280	.56	.497	.007

One-Sample Test

	Test Value = .583333					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
TotalFalse	-3.520	5279	.00044	-.024	-.04	-.01

APPENDIX VV.

SPSS output, Study Two, self-report ratio descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
None_SR_ratio	106	.00	.43	.1455	.08744
Listened_SR_ratio	106	.00	.50	.3374	.07602
Watched_SR_ratio	106	.14	1.00	.3437	.09515
Both_SR_ratio	106	.00	.25	.1731	.05445
Valid N (listwise)	106				

APPENDIX WW.

SPSS output, Study Two, Pearson's correlation

analysis for individual differences for self-report ratio scores

		AEI Experience	AEI Special Experience	AEI Belief	AEI Special Belief
None_SR_ratio	Pearson Correlation	-.236	-.215	-.051	-.032
	Sig. (2-tailed)	.236	.281	.802	.873
	N	27	27	27	27
Listened_SR_ratio	Pearson Correlation	-.232	-.234	-.146	-.159
	Sig. (2-tailed)	.244	.240	.466	.428
	N	27	27	27	27
Watched_SR_ratio	Pearson Correlation	.410*	.406*	.263	.261
	Sig. (2-tailed)	.034	.036	.185	.188
	N	27	27	27	27
Both_SR_ratio	Pearson Correlation	-.315	-.326	-.308	-.308
	Sig. (2-tailed)	.109	.097	.118	.118
	N	27	27	27	27

		Unusual Experiences	Cognitive Disorganisation	Introvertive Anhedonia	Impulsive Nonconformity	CPES
None_SR ratio	Pearson Correlation	.105	-.010	.128	.330	.352
	Sig. (2-tailed)	.602	.961	.526	.100	.071
	N	27	26	27	26	27
Listened_SR ratio	Pearson Correlation	-.162	-.174	.175	-.008	.003
	Sig. (2-tailed)	.419	.396	.382	.969	.990
	N	27	26	27	26	27
Watched_SR ratio	Pearson Correlation	.063	.145	-.242	-.297	-.165
	Sig. (2-tailed)	.755	.480	.223	.140	.412
	N	27	26	27	26	27
Both_SR ratio	Pearson Correlation	.015	-.068	.138	.357	-.004
	Sig. (2-tailed)	.942	.741	.493	.074	.984
	N	27	26	27	26	27

APPENDIX XX.

SPSS output, Study Two, self-report ratio regression analysis

Regression analysis 1:

Predictor Variables: 9 x Individual differences measures

Outcome Variable: Self-report ratio score – 'none' condition

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.536[a]	.288	-.113	.07885

a. Predictors: (Constant), CPES_Score, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Introvertive_Anhedonia_Total, AEI_experience_mean, O_LIFE_Unusual_Experiences_Total, AEI_Belief_Score_inc_last_two_scores_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	.040	9	.004	.718	.687[b]
	Residual	.099	16	.006		
	Total	.140	25			

a. Dependent Variable: None_SR_ratio

b. Predictors: (Constant), CPES_Score, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Introvertive_Anhedonia_Total, AEI_experience_mean, O_LIFE_Unusual_Experiences_Total, AEI_Belief_Score_inc_last_two_scores_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.044	.122		.366	.719
	AEI_experience_mean	-.386	.855	-2.055	-.452	.657
	AEI_Experience_Score_inc_last_two_scores_mean	.309	.915	1.601	.337	.740
	AEI_belief_mean	.051	.383	.284	.134	.895
	AEI_Belief_Score_inc_last_two_scores_mean	.007	.413	.037	.017	.986
	O_LIFE_Unusual_Experiences_Total	-.002	.015	-.038	-.117	.908
	O_LIFE_Cognitive_Disorganisation_Total	-.001	.010	-.036	-.137	.893
	O_LIFE_Introvertive_Anhedonia_Total	-.003	.013	-.055	-.223	.827
	O_LIFE_Impulsive_Nonconformity_Total	.010	.010	.257	1.027	.320
	CPES_Score	.007	.007	.285	.998	.333

a. Dependent Variable: None_SR_ratio

Regression analysis 2:

Predictor Variables: 9 x Individual differences measures

Outcome Variable: Self-report ratio score– ‘listened to’ condition

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.508[a]	.258	-.160	.10959

a. Predictors: (Constant), CPES_Score, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Introvertive_Anhedonia_Total, AEI_experience_mean, O_LIFE_Unusual_Experiences_Total, AEI_Belief_Score_inc_last_two_scores_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	.067	9	.007	.618	.766[b]
	Residual	.192	16	.012		
	Total	.259	25			

a. Dependent Variable: Listened_SR_ratio

b. Predictors: (Constant), CPES_Score, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Introvertive_Anhedonia_Total, AEI_experience_mean, O_LIFE_Unusual_Experiences_Total, AEI_Belief_Score_inc_last_two_scores_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	T	Sig.
1	(Constant)	.479	.169		2.833	.012
	AEI_experience_mean	.244	1.188	.951	.205	.840
	AEI_Experience_Score_inc_last_two_scores_mean	-.338	1.272	-1.289	-.266	.794
	AEI_belief_mean	.535	.532	2.177	1.005	.330
	AEI_Belief_Score_inc_last_two_scores_mean	-.527	.574	-1.970	-.917	.373
	O_LIFE_Unusual_Experiences_Total	-.027	.021	-.435	-1.322	.205

	O_LIFE_Cognitive_Disorganisation_Total	-.013	.013	-.264	-.971	.346
	O_LIFE_Introvertive_Anhedonia_Total	.022	.018	.310	1.237	.234
	O_LIFE_Impulsive_Nonconformity_Total	.004	.014	.081	.317	.755
	CPES_Score	.005	.010	.156	.535	.600

a. Dependent Variable: Listened_SR_ratio

Regression analysis 3:

Predictor Variables: 9 x Individual differences measures

Outcome Variable: Self-report ratio score– ‘watched’ condition

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.653[a]	.427	.104	.14789

a. Predictors: (Constant), CPES_Score, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Introvertive_Anhedonia_Total, AEI_experience_mean, O_LIFE_Unusual_Experiences_Total, AEI_Belief_Score_inc_last_two_scores_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	.261	9	.029	1.324	.299[b]
	Residual	.350	16	.022		
	Total	.610	25			

a. Dependent Variable: Watched_SR_ratio

b. Predictors: (Constant), CPES_Score, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Introvertive_Anhedonia_Total, AEI_experience_mean, O_LIFE_Unusual_Experiences_Total, AEI_Belief_Score_inc_last_two_scores_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.398	.228		1.743	.101
	AEI_experience_mean	-.949	1.603	-2.415	-.592	.562
	AEI_Experience_Score_inc_last_two_scores_mean	1.222	1.716	3.031	.712	.487
	AEI_belief_mean	-.439	.718	-1.162	-.611	.550
	AEI_Belief_Score_inc_last_two_scores_mean	.362	.775	.883	.468	.646
	O_LIFE_Unusual_Experiences_Total	.027	.028	.278	.959	.352
	O_LIFE_Cognitive_Disorganisation_Total	.015	.018	.200	.837	.415
	O_LIFE_Introvertive_Anhedonia_Total	-.029	.024	-.263	-1.195	.250
	O_LIFE_Impulsive_Nonconformity_Total	-.023	.018	-.282	-1.256	.227
	CPES_Score	-.018	.014	-.339	-1.325	.204

a. Dependent Variable: Watched_SR_ratio

Regression analysis 4:

Predictor Variables: 9 x Individual differences measures

Outcome Variable: Self-report ratio score– ‘both’ condition

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.619[a]	.383	.036	.06492

a. Predictors: (Constant), CPES_Score, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Introvertive_Anhedonia_Total, AEI_experience_mean, O_LIFE_Unusual_Experiences_Total, AEI_Belief_Score_inc_last_two_scores_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	.042	9	.005	1.103	.413[b]
	Residual	.067	16	.004		
	Total	.109	25			

a. Dependent Variable: Both_SR_ratio

b. Predictors: (Constant), CPES_Score, O_LIFE_Cognitive_Disorganisation_Total, O_LIFE_Impulsive_Nonconformity_Total, O_LIFE_Introvertive_Anhedonia_Total, AEI_experience_mean, O_LIFE_Unusual_Experiences_Total, AEI_Belief_Score_inc_last_two_scores_mean, AEI_belief_mean, AEI_Experience_Score_inc_last_two_scores_mean

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	T	Sig.
1	(Constant)	.069	.100		.688	.501
	AEI_experience_mean	1.078	.704	6.482	1.531	.145
	AEI_Experience_Score_inc_last_two_scores_mean	-1.184	.753	-6.944	-1.572	.136
	AEI_belief_mean	-.140	.315	-.876	-.444	.663

	AEI_Belief_Score_inc_last_two_scores_mean	.153	.340	.884	.451	.658
	O_LIFE_Unusual_Experiences_Total	.004	.012	.095	.317	.755
	O_LIFE_Cognitive_Disorganisation_Total	.000	.008	-.008	-.031	.975
	O_LIFE_Introvertive_Anhedonia_Total	.008	.011	.174	.762	.457
	O_LIFE_Impulsive_Nonconformity_Total	.009	.008	.253	1.085	.294
	CPES_Score	.006	.006	.263	.990	.337

a. Dependent Variable: Both_SR_ratio

APPENDIX YY.

SPSS output, Study Three, overall participant descriptive statistics

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
iValid	Male	34	34.0	34.0	34.0
	Female	66	66.0	66.0	100.0
	Total	100	100.0	100.0	

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age	100	18	67	35.87	13.078
Valid N (listwise)	100				

APPENDIX ZZ.

SPSS output, Study Three, ‘safe’ area participant descriptive statistics

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	29	39.2	39.2	39.2
	Female	45	60.8	60.8	100.0
	Total	74	100.0	100.0	

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Gender	74	1	2	1.61	.492
Valid N (listwise)	74				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age	74	18	67	36.78	13.662
Valid N (listwise)	74				

APPENDIX AAA.

SPSS output, Study Three, ‘threatening’ area participant descriptive statistics

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	5	19.2	19.2	19.2
	Female	21	80.8	80.8	100.0
	Total	26	100.0	100.0	

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Gender	26	1	2	1.81	.402
Valid N (listwise)	26				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age	26	18	63	33.27	11.080
Valid N (listwise)	26				

APPENDIX BBB.

A photo of the ‘safe’ location

APPENDIX CCC.

A photo of the ‘threatening’ location

APPENDIX DDD.

Study Three Coding

To code each field study participant, use the below (descending from top to bottom) to create a reference.

Age	If the participant appears to be 20-29, the code starts 20, if the partcipant appears to be 30-39, the code starts 30 and so on.
Gender	Male = M, Female = F
Hair colour	Dark = D, Medium = M, Light = L
Hair length	Long = L, Medium = M, Short = S
Coat colour	Black = B, Brown = Br, Green = G, Blue = Bl, Red = R, Other = O
Example 1	A woman with long brown hair in a blue coat who appears to be in her 30s would be coded: **30FMLBl**
Example 2	A man who appears to be in his 20s with medium length blonde hair who wore a brown coat would be coded: **20MLMBr**

APPENDIX EEE.

SPSS output, Study Three, how often participants 'felt' surveillance in the 'safe' area

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
FeltSurveillance	74	0	1	.51	.503
Valid N (listwise)	74				

APPENDIX FFF.

SPSS output, Study Three, how often participants 'felt' surveillance in the 'threatening' area

	N	Minimum	Maximum	Mean	Std. Deviation
FeltSurveillance	26	0	1	.58	.504
Valid N (listwise)	26				

APPENDIX GGG

SPSS output, Study Three, how often participants 'felt' surveillance in the overall areas

	N	Minimum	Maximum	Mean	Std. Deviation
FeltSurveillance	100	0	1	.53	.502
Valid N (listwise)	100				

APPENDIX HHH.

SPSS output, Study Three, correct mean responses in the 'safe' area

	N	Minimum	Maximum	Mean	Std. Deviation
Correct	74	0	1	.50	.503
Correct_Surveillance	74	0	1	.26	.440
Correct_Non_Surveillance	74	0	1	.24	.432
Valid N (listwise)	74				

APPENDIX III.

SPSS output, Study Three, correct mean responses in the 'threatening' area

	N	Minimum	Maximum	Mean	Std. Deviation
Correct	26	0	1	.69	.471
Correct_Surveillance	26	0	1	.38	.496
Correct_Non_Surveillance	26	0	1	.31	.471
Valid N (listwise)	26				

APPENDIX JJJ.

SPSS output, Study Three, binomial test for the 'safe' area

Hypothesis Test Summary

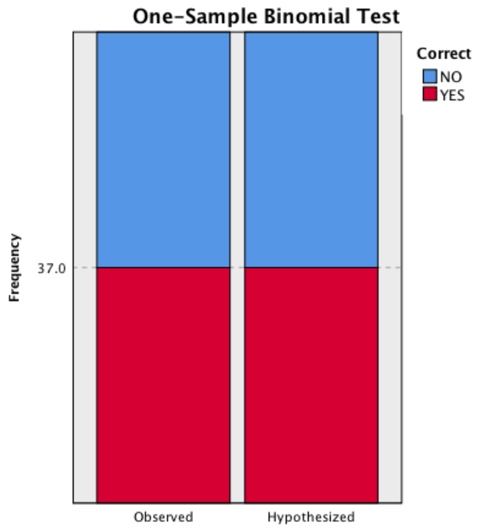
	Null Hypothesis	Test	Sig.	Decision
1	The categories defined by Correct = NO and YES occur with probabilities .500 and .500.	One-Sample Binomial Test	1.000	Retain the null hypothesis.

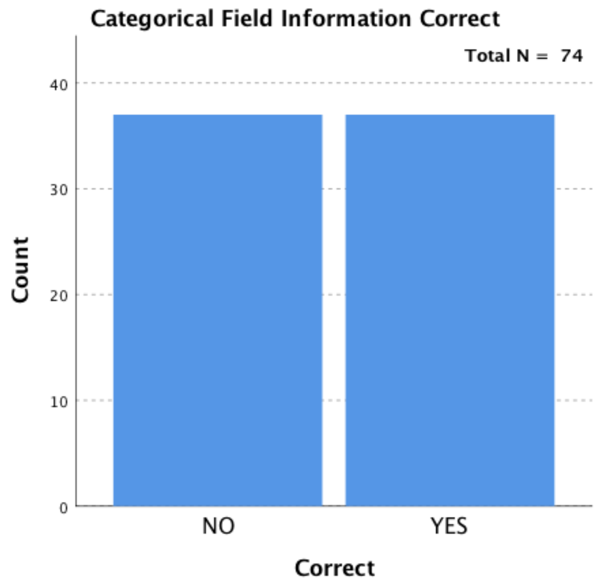
Confidence Interval Summary

Confidence Interval Type	Parameter	Estimate	95.0% Confidence Interval	
			Lower	Upper
One-Sample Binomial Success Rate (Clopper-Pearson)	Probability(Correct=NO).	.500	.381	.619

One-Sample Binomial Test Summary

Total N	74
Test Statistic	37.000
Standard Error	4.301
Standardized Test Statistic	.000
Asymptotic Sig.(2-sided test)	1.000





APPENDIX KKK.

SPSS output, Study Three, binomial test for the 'threatening' area

Hypothesis Test Summary

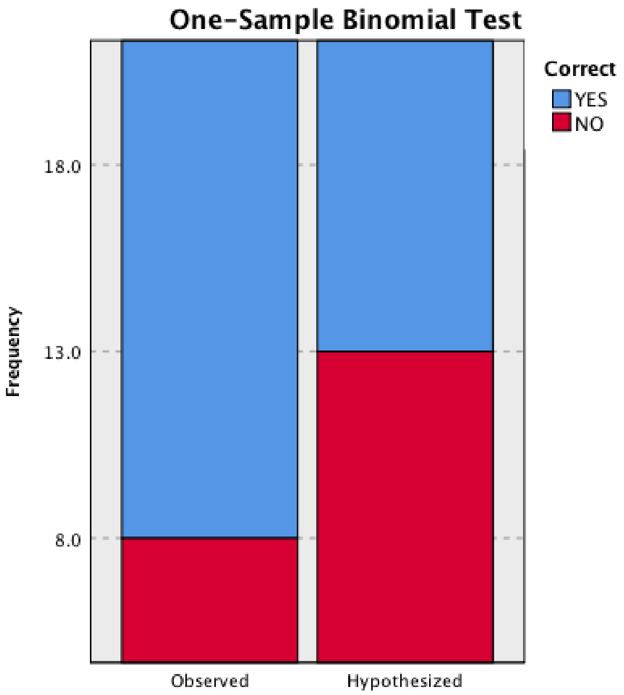
	Null Hypothesis	Test	Sig.	Decision
1	The categories defined by Correct = YES and NO occur with probabilities .500 and .500.	One-Sample Binomial Test	.078	Retain the null hypothesis.

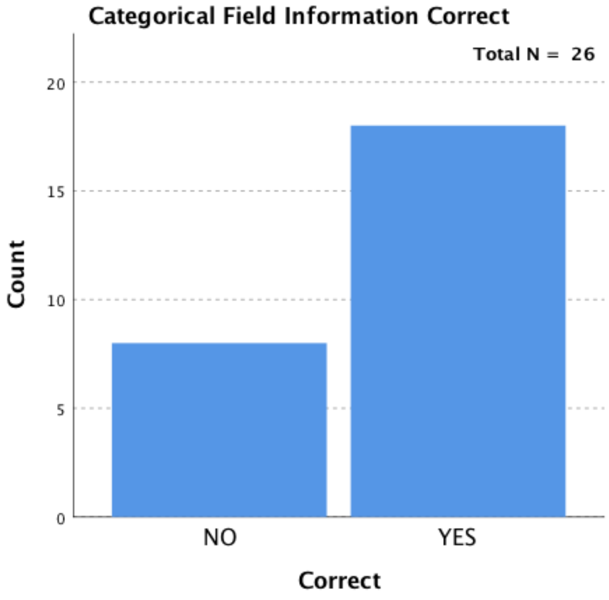
Confidence Interval Summary

Confidence Interval Type	Parameter	Estimate	95.0% Confidence Interval Lower	Upper
One-Sample Binomial Success Rate (Clopper-Pearson)	Probability(Correct=YES).	.692	.482	.857

One-Sample Binomial Test Summary

Total N	26
Test Statistic	18.000
Standard Error	2.550
Standardized Test Statistic	1.765
Asymptotic Sig.(2-sided test)	.078





APPENDIX LLL.

SPSS output, Study Three, Chi-Square Tests for Comparsisons Between Areas.

Correct * Area Crosstabulation

Count

		Area		
		SAFE	DANGEROUS	Total
Correct	NO	37	8	45
	YES	37	18	55
Total		74	26	100

Chi-Square Tests

	Value	Df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.875[a]	1	.090		
Continuity Correction[b]	2.150	1	.143		
Likelihood Ratio	2.945	1	.086		

Fisher's Exact Test				.111	.070
Linear-by-Linear Association	2.846	1	.092		
N of Valid Cases	100				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 11.70.
b. Computed only for a 2x2 table

APPENDIX MMM.

SPSS output, Study Three, Chi-Square Tests for Comparsisons Within Areas.

ActualSurveillance * Correct Crosstabulation

Count

		Correct		Total
		NO	YES	
ActualSurveillance	NO	19	18	37
	YES	18	19	37
Total		37	37	74

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.054[a]	1	.816		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.054	1	.816		
Fisher's Exact Test				1.000	.500
Linear-by-Linear Association	.053	1	.817		
N of Valid Cases	74				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 18.50.
b. Computed only for a 2x2 table

ActualSurveillance * Correct Crosstabulation

Count

		Correct		Total
		NO	YES	
ActualSurveillance	NO	5	8	13
	YES	3	10	13
Total		8	18	26

Chi-Square Tests

	Value	Df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.722[a]	1	.395		
Continuity Correction[b]	.181	1	.671		
Likelihood Ratio	.728	1	.394		
Fisher's Exact Test				.673	.336
Linear-by-Linear Association	.694	1	.405		
N of Valid Cases	26				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 4.00.

b. Computed only for a 2x2 table

APPENDIX NNN.

SPSS output, Study Three, Individual Differences descriptive statistics for the ‘threatening’ area

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Correct_Surveillance	26	0	1	.38	.496
Correct_Non_Surveillance	26	0	1	.31	.471
Light	26	9000	15000	11384.62	1395.046
Speed	26	23	40	29.69	4.343

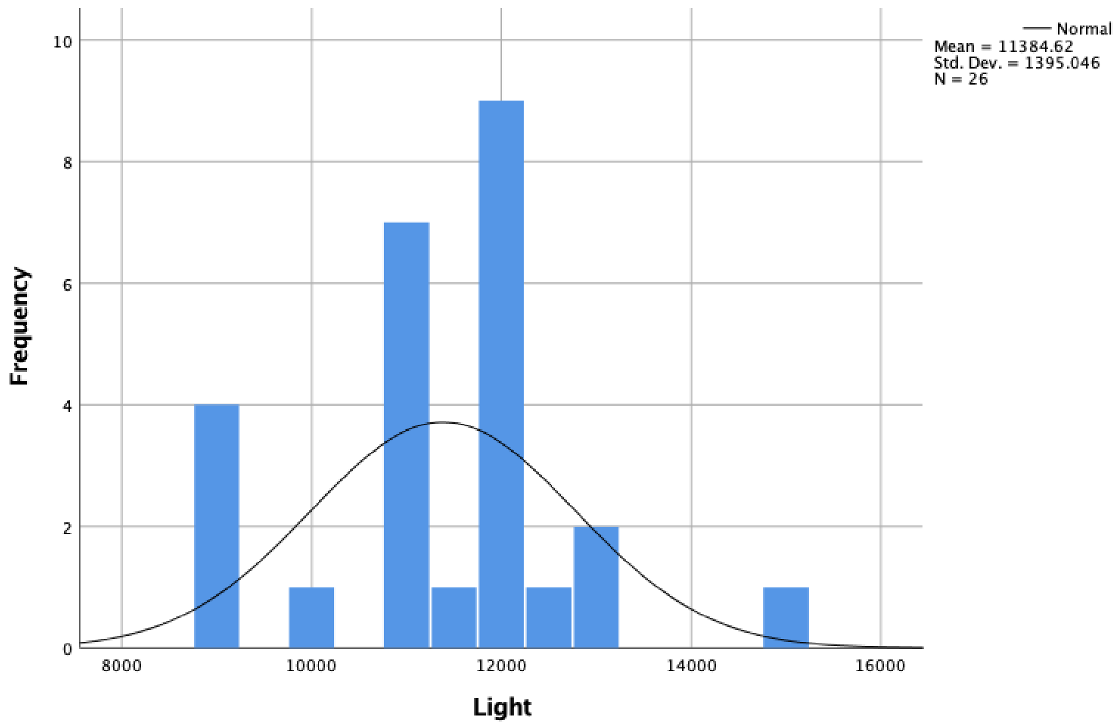
No_Of_Head_Turns	26	0	3	.50	.762
Degree_Of_Head_Turns	26	0	3	.62	.898
No_Of_Adjustments	26	0	3	.54	.859
No_Of_Pauses	26	0	0	.00	.000
Valid N (listwise)	26				

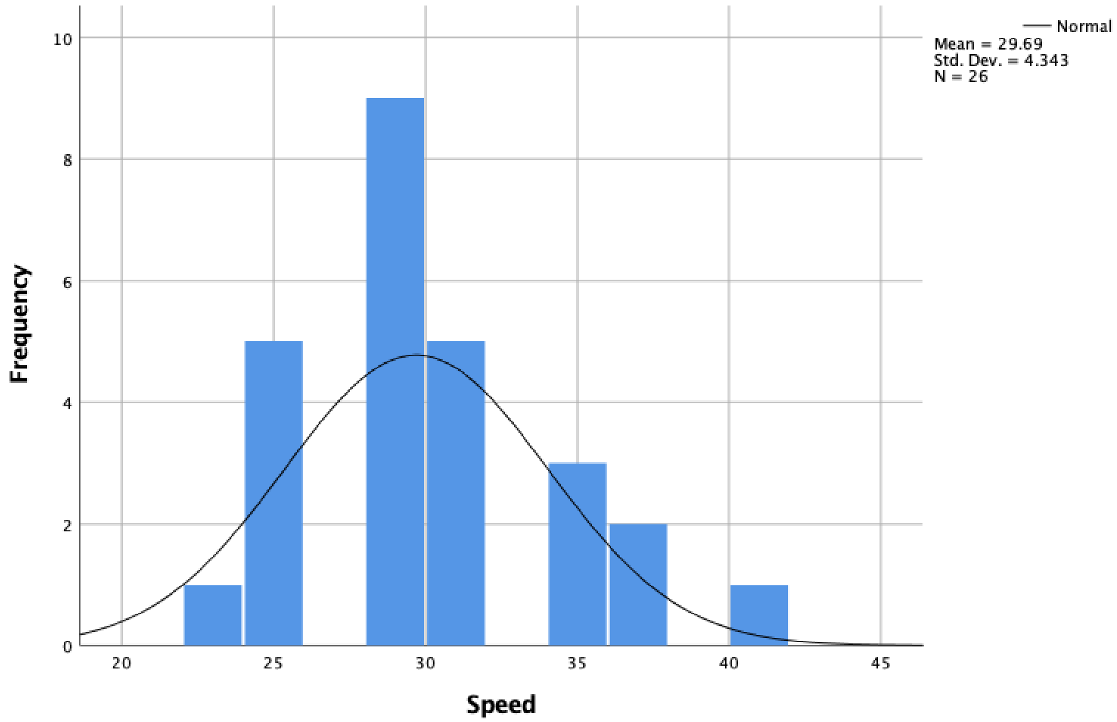
APPENDIX OOO.

SPSS output, Study Three, passing the assumptions for t-tests

Group Statistics

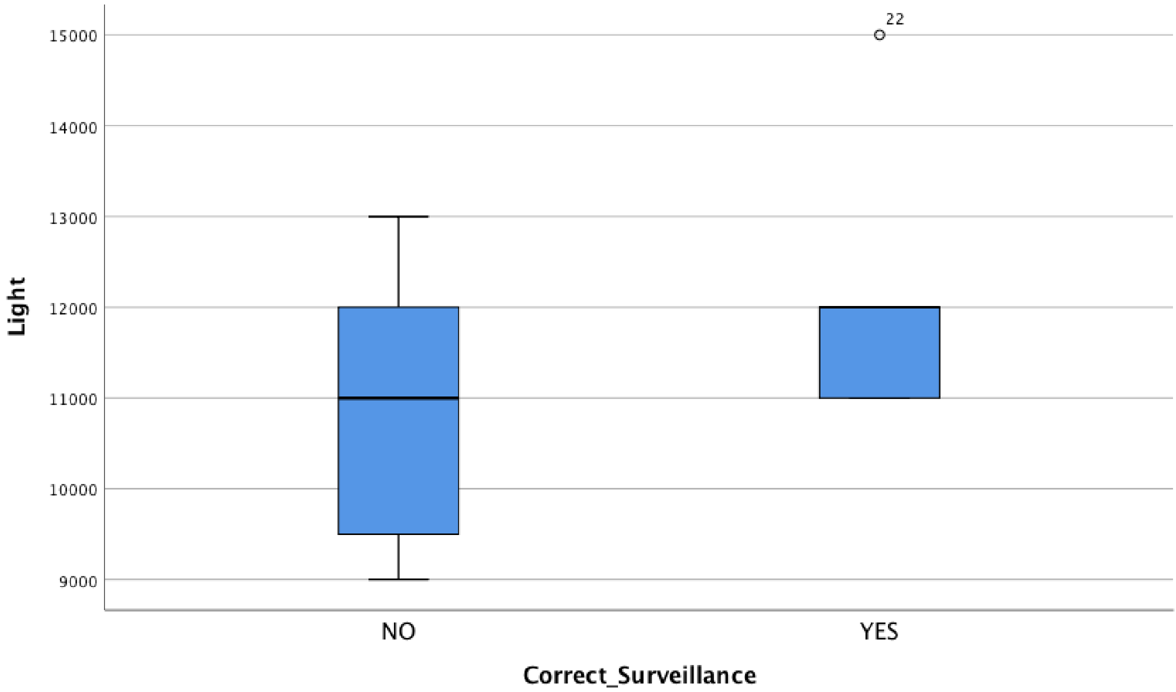
	Correct	N	Mean	Std. Deviation	Std. Error Mean
Light	NO	45	11088.89	1411.354	210.392
	YES	55	11327.27	1401.779	189.016
Speed	NO	45	28.44	4.635	.691
	YES	55	29.87	4.363	.588
No_Of_Head_Turns	NO	37	.70	1.077	.177
	YES	37	.65	1.006	.165
Degree_Of_Head_Turns	NO	37	.73	1.045	.172
	YES	37	.62	.861	.142


Normal
Mean = 11384.62
Std. Dev. = 1395.046
N = 26
Frequency
Light
0
2
4
6
8
10
8000
10000
12000
14000
16000


Normal
Mean = 29.69
Std. Dev. = 4.343
N = 26
Frequency
Speed
0
2
4
6
8
10
20
25
30
35
40
45

Tests of Normality

	Correct_Surveillance	Kolmogorov-Smirnov[a] Statistic	Df	Sig.	Shapiro-Wilk Statistic	df	Sig.
Light	NO	.188	16	.136	.894	16	.065
	YES	.400	10	.000	.671	10	.000

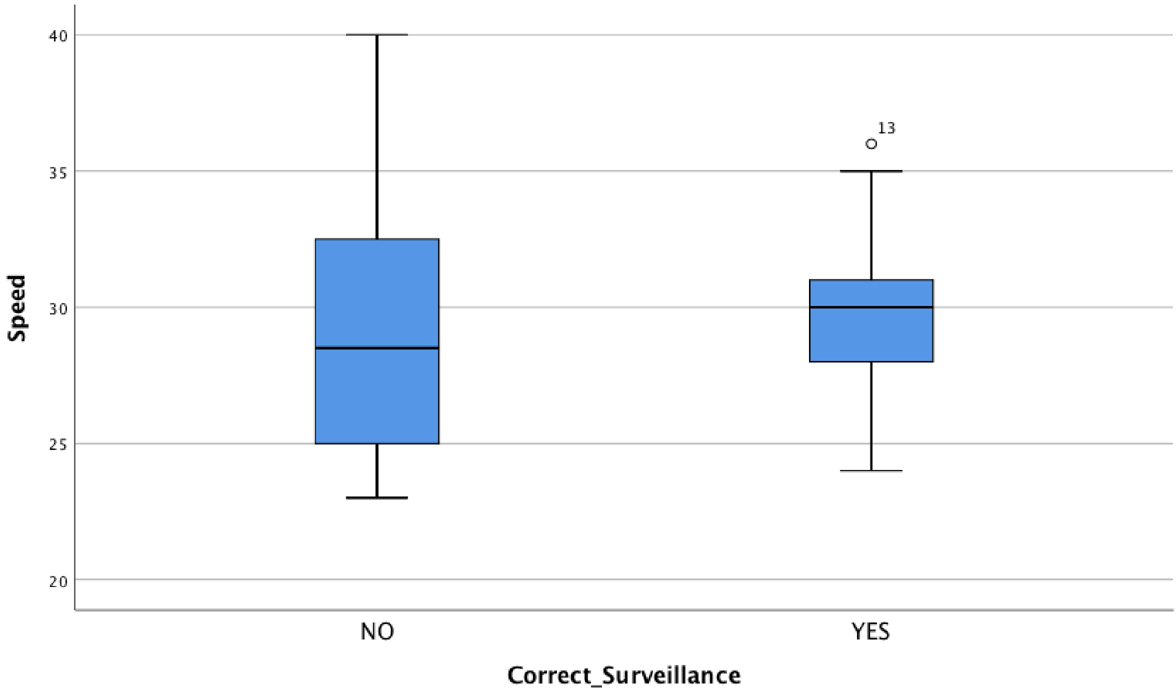


Tests of Normality

	Correct_Surveillance	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
Speed	NO	.223	16	.032	.896	16	.069
	YES	.199	10	.200*	.938	10	.530

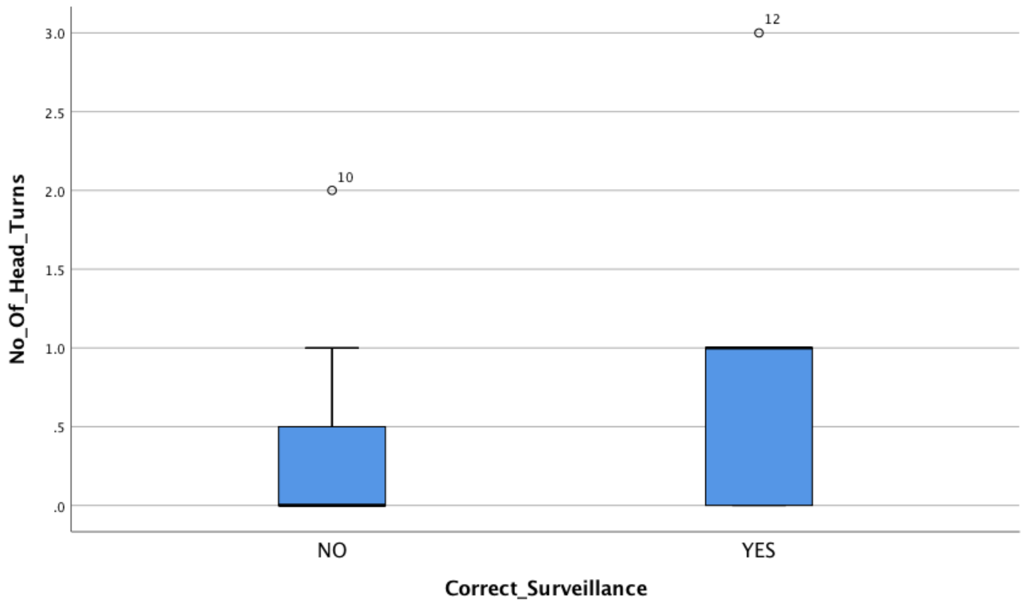
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction



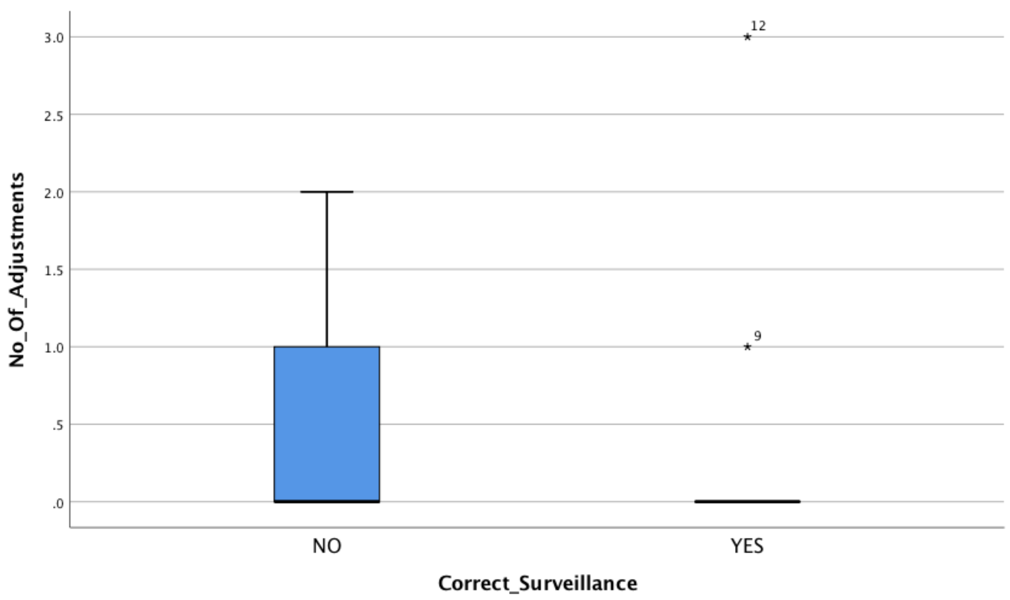
Tests of Normality

	Correct_Surveillance	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	df	Sig.
No_Of_Head_Turns	NO	.448	16	.000	.587	16	.000
	YES	.314	10	.006	.750	10	.004



Tests of Normality

	Correct_Surveillance	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	df	Sig.
No_Of_Adjustments	NO	.343	16	.000	.732	16	.000
	YES	.461	10	.000	.500	10	.000



APPENDIX PPP.

SPSS output, Study Three, Independent T-Tests on Environmental and Behavioural Differences

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Light	Equal variances assumed	5.105	.033	-.625	24	.538	-375.000	600.142	-1613.633	863.633
	Equal variances not assumed			-.536	9.917	.604	-375.000	700.130	-1936.762	1186.762
Speed	Equal variances assumed	.189	.668	-1.907	24	.069	-3.347	1.755	-6.970	.275
	Equal variances not assumed			-2.002	15.183	.063	-3.347	1.672	-6.907	.213
No_Of_Head_Turns	Equal variances assumed	5.937	.023	-1.740	24	.095	-.542	.311	-1.184	.101
	Equal variances not assumed			-2.313	23.995	.030	-.542	.234	-1.025	-.058
Degree_Of_Head_Turns	Equal variances assumed	11.176	.003	-1.959	24	.062	-.708	.362	-1.455	.038
	Equal variances not assumed			-2.686	23.490	.013	-.708	.264	-1.253	-.163

APPENDIX QQQ.

SPSS output, Study Three, Chi-Square Test on Behavioural Differences

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Correct * NominalHeadturn	26	100.0%	0	0.0%	26	100.0%

Correct * NominalHeadturn Crosstabulation

Count

		NominalHeadturn		
		0	1	Total
Correct	NO	7	1	8
	YES	9	9	18
Total		16	10	26

Chi-Square Tests

	Value	Df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.291[a]	1	.070		
Continuity Correction[b]	1.897	1	.168		
Likelihood Ratio	3.665	1	.056		
Fisher's Exact Test				.099	.081
Linear-by-Linear Association	3.164	1	.075		
N of Valid Cases	26				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 3.08.

b. Computed only for a 2x2 table

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Correct * NominalClothing	26	100.0%	0	0.0%	26	100.0%

Correct * NominalClothing Crosstabulation

Count

		NominalClothing		
		0	1	Total
Correct	NO	4	4	8
	YES	13	5	18
Total		17	9	26

Chi-Square Tests

	Value	Df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.208[a]	1	.272		
Continuity Correction[b]	.426	1	.514		
Likelihood Ratio	1.181	1	.277		
Fisher's Exact Test				.382	.255
Linear-by-Linear Association	1.162	1	.281		
N of Valid Cases	26				

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 2.77.

b. Computed only for a 2x2 table