Cognitive Bias and Forensic Anthropology: The Power of Context in the Interpretation of Skeletal Remains

Sherry Nakhaeizadeh

University College London

Ph.D. Forensic Science

April 2017

Declaration

I, Sherry Nakhaeizadeh confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed ____________________

Date ____________________

In loving memory of my dearest stepmom Maryam, who started this journey with me. Her love, support, and tremendous courage made me feel that anything is possible.

I carry you in my heart always. I hope I made you proud.

Acknowledgments

This research was funded by the Engineering and Physical Sciences Research Council of the U.K. (EPSRC), through the Security Science Doctoral Research Training Centre (UCL SECReT) based at University College London (EP/G037264/1).

I wish to thank and express my appreciation to my supervisors Dr Ruth Morgan, Dr Itiel Dror, and Dr Carolyn Rando for their support, encouragement, and expertise throughout the process. This thesis was made possible through their guidance and academic input.

For the love and support provided by my family, Mum, Melinda, Sanaz, Sahar, Sepideh: Thank you for always believing in me, I will always be grateful.

Special thanks are extended to Nathalie Dozzi for her devoted friendship and encouragement throughout this process, Nadine Smit, for her guidance, and our endless thought-provoking discussions, and to all the participants that took their personal time to participate in my experimental studies.

Last, but certainly not least, Tom thank you for your patience, love, and unconditional support.

Abstract

The central role of human cognition in forensic science and its effect on the interpretation of forensic evidence is being increasingly recognised within the forensic disciplines. It is clear that the concerns over expert decision-making and their vulnerabilities have not only been highlighted in recent key governmental reports, but also created a debate within the literature. This has, within recent years lead to a rise of empirical research focusing on the impact of cognitive biases in all stages of the forensic science process, highlighting that these vulnerabilities are not limited to a specific area of expertise.

In forensic anthropology, the presence of cognitive bias, its impact, and how to mitigate its effects are still not fully empirically assessed or appreciated. This thesis seeks to unearth and understand the degree to which contextual biases are present in forensic anthropology, and present ways that can mitigate the impacts in biological profiling. This research addresses the effect of context within forensic anthropological analysis throughout the forensic science process (collection, analysis, interpretation,) through a series of experimental studies. The results of the experimental studies showed that context could have a powerful effect in visual assessments of skeletal remains in sex, ancestry and age at death. Furthermore, the findings also provided an important first step towards understanding the potential effects of initial exposure to irrelevant context at a crime scene in the excavations of skeletal remains, showing a potential for cascading bias on the subsequent assessment of the skeletal remains.

An evidence-based approach for dealing with cognitive interpretation issues within the human identification field is presented. The findings of this thesis have contributed to the body of knowledge and provide empirical data that illustrate the benefits of developing a more holistic approach to forensic decision-making from crime scene to court within forensic anthropology and the wider forensic disciplines.

Table of Content

List of Tables

List of Figures

Chapter 1. Introduction

1.1. Overview

Forensic science has played an important role in criminal investigations and the legal process for centuries (Found 2014). In recent years, however, concerns with regard to the lack of rigorous scientific research within the forensic science domains have been expressed in the published literature and key governmental reports (National Acadamy of Science 2009; Government Chief Scientific Adviser 2015; Tully 2015; President's Committee of Advisors on Science and Technology 2016). The emergence of the recognition of cognitive bias within forensic science and criminal investigations is being increasingly discussed and described as an issue and concern in relation to the admissibility of evidence and expert witness testimony (Nakhaeizadeh et al. 2015). The research within the field of decision-making has emphasised the dynamic and active nature of human information processing and how it can lead to the distortion of incoming data, resulting in biased conclusions. The research within human cognition has shifted its focus to not only concern human judgments in the social and psychological domains, but has also emerged within law enforcement agencies and forensic disciplines (Kassin et al. 2013; Dror 2015).

The impact of cognitive biases has begun to be evaluated at all stages of the forensic science process including data collection, analysis, evidence interpretation and final presentation in court (Edmond et al. 2016; Found 2014). It has been demonstrated that these vulnerabilities are not limited to a specific field, with similar cognitive biasing issues being established across numerous forensic science domains (Dror et al. 2006; Dror & Hampikian 2011; Page et al. 2012; Osborne et al. 2014; Osborne et al. 2016; Stoel et al. 2014; Miller 1984; Klales & Lesciotto 2016; Nakhaeizadeh, Dror et al. 2014). Human decision-making is a key component of the forensic science process and has been shown to influence investigative decisions as well as legal outcomes, with the potential for significant societal impact at a global scale (Kassin et al. 2013). It is clear that the concerns raised over expert decision-making, including vulnerabilities of cognitive processes and inappropriate weight assigned to evidence, have created debate and heated controversy. Many have questioned the role of the forensic scientist at crime scenes and the exposure to contextual information, potentially being one of the sources to constitute bias in forensic settings. Specific criticism has been directed at the field of comparison and identification, including the specialisms of bite-mark comparison (Osborne et al. 2014), morphological hair analysis (Federal Bureau of Investigation 2015), fingerprints (Dror et al. 2011; Earwaker et al. 2015) and the use of ear prints (Champod et al. 2001) as evidence, as well as the field of forensic anthropology (Nakhaeizadeh & Morgan 2015).

Recently, there has been an increase in the critique of some of the techniques used by forensic anthropologists. Discussion has been extensive concerning evidence validation, admissibility, and error rates in the methods applied (Christensen & Crowder 2009; Christensen 2004). While many of the issues that have been identified have been addressed in new research within the field, some areas are yet to be fully addressed. The presence of cognitive bias, its impact, and the cognitive processes involved in the assessment of human remains have only recently begun to be assessed (Nakhaeizadeh, Dror, et al. 2014; Nakhaeizadeh, Hanson, et al. 2014; Klales & Lesciotto 2016).

It has been argued that some of the techniques used in forensic anthropology are generally reliant upon observation and the specialised experience of the observer (Cattaneo 2007; Byers 2010; Dirkmaat et al. 2008; Hefner et al. 2007), with some contesting the techniques and asserting that they are limited because of their subjective nature (Walrath et al. 2004). In response, there has been some modification of existing methods that have accompanied the development of new comparative samples in forensic anthropology and tools for data analysis (Grivas & Komar 2008) . These developments have enhanced the role of quantitative methods and have led to a rise in new publications in the literature. However, there is still considered to be variation and inconsistency among practitioners in the way in which methods are employed and how the results are reported when establishing a biological profile (SWGANTH 2016). Most quantitative techniques have still not been widely adopted, with many anthropologists preferring to use the traditional non-metric observation of morphological traits in sex, ancestry, and age at death estimations, all of which could be susceptible to cognitive interpretation issues.

The challenge of combining and interpreting different sources of information, and achieving transparency in decision-making and evidence-based conclusions needs to be tackled. Within the field of forensic anthropology, this includes a better understanding of the underlying processes of the decisions being made and potential cognitive influences in the interpretation of skeletal remains.

1.2. Aim and Research Questions

Therefore, the aim of this thesis was to further examine the extent to which cognitive biases are present within forensic anthropological methods. More specifically, this thesis sought to understand the degree of contextual effects in forensic anthropological assessments and thereby identify the means to avoid potential cognitive biases that may arise from interpretation issues.

In order to achieve this aim, the thesis sought to answer three main research questions:

1. Does contextual information such as grave context, and osteological reports affect and influence the interpretation process of visual assessments on skeletal remains on previously assessed skeletons?

2. Does initial exposure to 'extraneous' contexts at the crime scene in the excavations of skeletal remains affect upon judgments and interpretations in the subsequent skeletal analysis?

3. Does the order in which skeletal remains are assessed a) influences the interpretation of the subsequent skeletal element, (i.e. if examining a clear male pelvis will consequently skew the interpretation of the skull morphology and vice versa) and b) act as an influence and determine the final conclusion of the assessment?

In order to answer the research questions, a series of experiments were undertaken to test for cognitive and contextual effects empirically within forensic anthropological methods and procedures. More specifically, this was done using visual methods used in the establishment of a biological profile, focusing on sex, ancestry and age at death, and varying the contextual information (research question 1). To address research question 2, the thesis examined the potential effects of initial exposure to 'extraneous' context, by providing participants with context at a crime scene before the subsequent assessment of the skeletal remains by the participants. Lastly, research question 3 was addressed by comparing the analyses of the skeletal remains and their relation with the order in which the skeletal remains were assessed. These objectives allowed for a holistic examination of the stages and methodological procedures when and to what extent cognitive factors may affect performances and render the judgements of participants to be compromised, and equally when they do not. This also allowed for identifying the means to avoid potential cognitive biases that might arise from interpretation issues in forensic anthropology.

1.3. Structure of the thesis

This thesis is structured as outlined below, and contains a literature review, three experimental chapters, a general discussion and a final summary conclusion.

Chapter 2 provides an overview of the pertinent published literature. As this thesis is a multidisciplinary research project combining several fields, the literature review covers the three major disciplines starting with an overview of human cognition. The human cognition section covers broad examples of research and theories within cognitive biases and how research within this area has been shown to affect decision-making. This is followed by examples of cognitive bias studies within the legal system. The second part of the literature review provides an overview of the forensic science domain and the conceptual framework in which forensic processes take place. This is followed by an insight into the role of forensic science in the legal system (drawing upon examples from the UK and U.S. mainly) and expert evidence. In addition, the role of human cognition and cognitive biases within the forensic science process was also covered, highlighting how these could affect expert performance and forensic decision-making, drawing on previous and current research, wrongful convictions, as well as proposed solutions on how to mitigate these effects. Following this, the latter and final part of the literature review provides a broad overview of forensic anthropology and its methods used in the establishment of a biological profile. Further, the forensic anthropology section presents some current research within the field, and illustrates why this sub-field within forensic science may be prone to cognitive interpretation issues.

Chapter 3 presents the first two experimental studies conducted for this thesis, examining the effect of contextual information on judgment and decision-making in some of the traditional forensic anthropological visual methods used (addressing research question 1). This chapter is divided into two parts with part one covering experiment 1 (a pilot study) and part 2 covering experiment 2 (that builds and develops experiment 1). Both studies specifically focused on whether contextual information can affect previous judgments when assessing skeletal remains of an ambiguous nature. The two experimental studies were designed in order to look into the visual assessment of participants in sex, ancestry and age at death on skeletal elements conducted during three different phases, (phase 1: Baseline control, phase 2: Context, phase 3: Reliability control). This was done in order to gain insight into whether the decisions of participants were consistent regardless of contextual influences. The results addressed research question 1, indicating that context appears to affect previous judgments. These findings therefore contribute to the overall aim of the thesis to further understand contextual effects in biological profiling.

Chapter 4 presents the third experimental study that addresses the potential effects of initial exposure to context at a crime scene upon judgment and decision-making in the subsequent assessment of skeletal remains (addressing research question 2). This study specifically examined whether early exposure to 'extraneous' contexts in the excavation of skeletal remains cascade, and thereby affect the subsequent assessment of the skeletal analysis. This chapter primarily focused on contextual biases at the crime scene and their potential to lead to cascading affects. The specific research question addressed in this chapter was whether clothing associated with skeletal excavations could impact the evaluations and judgments of participants. The study was designed to investigate whether early exposure to such contexts could influence the primary working hypotheses. The findings indicated that early exposure to 'extraneous' context at the crime scene can affect subsequent assessments of the participants in the laboratory. The results further contribute to the overall aim of understanding contextual biases within different stages (excavation and interpretation) of skeletal analysis, in addition to what extent cognitive factors may affect performances and render the judgements of participants.

Chapter 5 presents the fourth experimental study presented in this thesis. This study exclusively focused on visual sex assessments within forensic anthropology. The analyses of the skeletal remains and the relationship of that analysis with the order in which the skeletal remains were assessed (addressing research question three) was examined. This research was designed to assess whether the order in which participants assessed skeletal remains for establishing a sex estimation, could influence the interpretation of the subsequent skeletal element, and/or act as an influence and thereby determine the final conclusion reached in the assessment. For example, the study sought to establish whether if a participant started a sex assessment by looking at a clear male pelvis would that observation consequently affect the interpretation of the skull morphology, and vice versa. This chapter presents data with regards to the degree of contextual effects in forensic anthropological methods that may arise from the order of examination and procedural practice. The results thereby contribute to the achievement of the overall aim of further understanding the degree of different types of influences that may affect interpretations in skeletal analyses.

Chapter 6 of this thesis presents an overarching discussion that draws together the key findings from all four experimental studies. The chapter presents the implications of the results as they relate to contextual biases in forensic anthropology specifically, and forensic science more broadly. Furthermore, the discussion in this chapter also highlights the importance of producing empirical data that can contribute to an evidence base that presents the extent to which cognitive factors may be influencing decision-making. The current debate regarding how to best test for and manage contextual biases throughout the forensic science process is outlined, and it is argued that empirical studies are the best means of identifying the best steps forward. Following this, the limitations of this thesis as well as future recommendation and directions for further future work are presented.

Chapter 7 provides the conclusion, which sets out the key findings in relation to each research question and presents the implications of this research for forensic anthropology specifically and the forensic sciences more broadly. The chapter concludes that context is influential in visual methods used in forensic anthropology (in the establishment of a biological profile) where human judgment plays a central role. Moreover, the power of contextual influences in the assessment of skeletal remains will differ and in some cases may result in contextual biases.

1.4. Additional considerations

It is important to highlight that due to the Human Tissue Act 2004 (c30) (applying to England, Northern Ireland and Wales) the practice of the use of modern skeletal remains in the experimental studies was constrained. Thus, all experimental studies in this thesis included skeletal remains from archaeological excavations. In addition, parts of the experimental research design included deception of participants. Therefore, appropriate ethical approval in accordance with UCL REC ethics committee (ethics nr 4672/001) was obtained and the data stored according to the Data Protection Act 1998.

Parts of the literature review for this thesis (section 2.1, 2.2, and 2.3) have previously been published in peer review journals by the author (see Appendix A). In addition, Chapter 3 (experimental study 2) and Chapter 4 (experimental study 3) are currently under review for publications (see Appendix A).

Chapter 2. Literature Review

2.1. Human Cognition

In order to appreciate how judgments and interpretations in forensic science and criminal investigations can be affected by cognitive mechanisms, it is important to recognise the strengths and weakness of human cognition in decision-making (Caverni et al. 1990). The approach of processing information in decision-making is known as human cognition, and defines the acquisition, organization and the use of knowledge (Anderson 2000; Wyer & Srull 1986; Bandura & Albert 1986). The study of human cognition examines human perception, judgment and decision-making, which are all influenced by a variety of cognitive processes (Hoppitt et al. 2010). The information-processing network in the human brain is very complex and in order for the brain to organise information it will use schemata to structure information and encode the relationship among them.

2.1.1. Schema and Cognitive Bias

Schemata play a vital role in judgment and decision-making, which generally are defined as "scripts" that help the brain analyse the perception and judgment of an individual based on their prior beliefs and experiences (Neisser 1976). The human mind encode the information coming in, which is known as 'bottom up' and is considered to be purely raw data derived from the environment. The processing and interpretation of incoming data (bottom-up information) is mediated by a variety of 'top down' cognitive mechanisms such as knowledge, experience, motivations expectations and emotional states (Kassin et al. 2013). Top-down processing makes the processing of information much more efficient (Dror & Kosslyn 1998) however, in some cases top-down components can interfere with and distort the processing of the bottom-up component (Fraser-Mackenzie et al. 2013) . For example, research within psychology and social science has demonstrated that the emotional state of individuals can have a significant impact upon the way information is processed and interpreted, as perceptions and understandings are highly related to emotional states (Byrne & Eysenck 1993). For instance, mock juror studies that have examined the issue of emotional state and decision-making have demonstrated that presenting emotionally disturbing evidence influences the verdict of mock jurors (Bright & Goodman-Delahunty 2006).

Within forensic science, it is now acknowledged that forensic case work can also potentially be influenced by a variety of top-down processing mechanisms, with much forensic analysis arguably occurring in highly emotional contexts where evidence is associated with specific crimes against a victim(s) (Dror et al. 2005). Therefore, relying on top down cognitive mechanisms and operative information processing is liable to cause susceptibility to weaknesses in the interpretation of evidence (Dror 2011). This type of information may affect the analytical methods and influence the decision-making procedure when generating the final conclusion and thereby cause a biasing effect (Girotto & Politzer 1990). This could be referred to as cognitive biases, generally defined as the psychological and cognitive factors that unconsciously manipulate and interfere with the data processing, causing judgment and decision-making to be unreliable (Evans & Pollard 1990). Cognitive biases are also part of a concept commonly known as heuristics.

2.1.2. Heuristics

Heuristics are strategies that use mental shortcuts in decision-making, including ignoring part of the information to make decisions quicker, more prudent and accurate (Gigerenzer & Gaissmaier 2011). For enhanced and frugal cognition, heuristics trade off some loss in accuracy, which could lead to faulty reasoning (Elstein 1999). The concept of heuristics was originally introduced by Simon's (1957), work of "Bounded Rationality". One of the main findings of Simon's (1957) research was that rationalities of individuals in decision-making is restricted by cognitive limitations, as people tend to accept judgments and choices that are satisfactory enough for their purpose (Simon 1957).

There are different scenarios in which heuristics may operate, such as through anchoring and adjustments, whereby the tendency is to rely on the first piece of information presented when making a decision (Bergman et al. 2010). For example studies regarding sentencing guidelines have demonstrated that judges use different judgmental anchors when making sentencing decisions (Englich & Musseweiler 2001). Judges were influenced by sentencing demands, which resulted in people who had committed very similar crimes receiving different sentences (Musseweiler & Englich 2005). Tversky & Kahneman (1975) demonstrated in their study that people tend to rely on various cognitive heuristics, and whilst this is considered generally to be beneficial (Gigerenzer & Gaissmaier 2011), it could also create systematic errors in judgment and decision-making. This has been specifically demonstrated when it comes to prior expectations, which could provide a sufficient and unconscious tendency to perceive and interpret evidence that would confirm pre-existing beliefs, otherwise known as confirmation bias (Khaneman & Frederick 2002).

2.1.3. Confirmation Bias

Confirmation bias is the tendency to selectively gather and process information to confirm a hypothesis or preconception (Dror & Charlton 2006) by looking for evidence that would validate existing beliefs and expectations, in terms of rejecting, excusing, or ignoring evidence that could contradict the current assumption (Gianelli 2007). Studies within reasoning have demonstrated that people attempt to find evidence, which confirms a hypothesis rather than finding evidence that would disconfirm it (Cheng et al. 1986). The fundamental mechanisms upon which confirmation bias operates are selective attention to information and biased interpretation of available information (Ask & Granhag 2005). Selective information search within legal contexts can be identified when an individual examines information or evidence to incriminate a suspect based on a personal hypothesis, and ignores evidence that could exonerate or lead to an alternative hypothesis.

Biased interpretations occur when experts only interpret evidence that supports, and will be in favour of their own hypotheses (Dror & Fraser-Mackenzie 2008). This inhibits the expert from observing the evidence from multiple angles, often resulting in a subjective conclusion (O'Brian 2009). For example, the majority of criminal investigations are driven by a theory, which leads investigators in their search for evidence to be guided by their initial hypothesis regarding when, why, how and by whom a crime was committed (Ask & Granhag 2005). These working hypotheses could arguably be affected by preconceptions and expectations of the investigators, due to the way the brain processes and stores information, especially when dealing with ambiguous and complex evidence (Burke 2005). Thus, a variety of influences that have nothing to do with the actual case drive can guide the investigation, and can affect its outcome. A preference for confirmation over falsification, could arguably result in investigators searching for and finding confirmatory evidence against a suspect in contrast to find disconfirming and exonerating information (O'Brian 2009).

The earliest work on confirmation biases can be traced back to the philosopher Francis Bacon who acknowledged its impact in his work of 1620 by recognizing various obstacles that influences the human mind (Dror 2009, Kassin et al. 2013). An important breakthrough and outcome of the study of reasoning and confirmation bias was the selection task of Wason (1966) with the findings of the classical experiment on card games that demonstrated that people attempt to find evidence which conforms to the rule rather than finding evidence that disconfirms it (Nickerson 1998; Sperber et al. 1995; Cheng et al. 1986). People prefer confirmation over falsification (Evans & Pollard 1990; Fiddick et al. 2000). This tendency to seek confirming evidence 'violated' (in that time) Karl Poppers prescription of "rational inference" where Popper's principle of rationality stated the need to seek falsification in testing of scientific hypotheses (Popper 1959).

Over the years confirmation bias has come to provide an umbrella term for a number of distinct ways that expectations and beliefs influence memory, selection, and evaluation of evidence (Nickerson 1998). For example, studies by Bruner and Potter (Bruner & Potter 1964), on interference in visual recognition demonstrated that expectations could have an impact upon perception. The study established that when participants were shown ambiguous images, they had the tendency to voluntarily generate a hypothesis about the vague images and then maintained these beliefs even as the real picture became clearer. Other phenomena also associated with confirmation bias are studies in belief persistence (Lord et al. 1979), overconfidence (Fischhoff et al. 1977), my side bias (Baron 1995), group conformity (Asch 1951) and memory and bias (Eagly et al. 1999).

2.1.4. Sources and Fuels of Confirmation Bias

The body of literature within psychology has over the years recognised different sources of cognitive bias, and confirmation bias in particular, such as time pressure (Ordóñez & Benson 1997), expectations (Bressan & Dal Martello 2002), pre-existing beliefs (Hamilton & Zanna 1974), and motivation (Kunda 1990). For example, a series of studies by Balcetis & Dunning (2006) showed that the impact of motivation on information processing lead participants to perceive a representation of the visual environment that they desired. Moreover, the studies demonstrated that participants tended to interpret an ambiguous figure in a manner that 'fitted' with their preference and wishes. This shows (together with decease of research in psychology) that perception is selective and malleable, and highly related to the context within which the decision is being made (Bugelski & Alampay 1961). For example, the understanding of how 'steep' a hill might be will be more extreme if participants are asked to make that estimation after they have jogged actively for an hour (Bhalla & Proffitt 1999). Similarly, an estimation of the speed of a person will be biased if participants are asked to make that estimation after viewing very fast animals (such as a cheetah) or very slow animals (such as a turtle) (Aarts & Dijksterhuis 2002). This highlights the fact that top down influences inform perceptions resulting in an impact on the perceptions of the human mind, which can result in the beliefs held by individuals being resistant to change. (Dror et al. 2005).

2.1.5. Belief Perseverance

Empirical research has demonstrated that once people form a hypothesis they can fail to adjust the tenacity of their beliefs in the light of evidence that will challenge the accuracy of those beliefs (Burke 2005). This is also known as belief perseverance, which is the tendency to continue to confirm to a theory even though the evidence underlying the theory is confounded (Anderson & Kellam 1992). One of the earliest studies in belief perseverance was to study the effect of what is known as the debriefing paradigm. In a study conducted by Anderson et al. (1980) subjects were presented with allegedly authentic reports of fire-fighters. After reading the reports subjects were asked to write an explanation of the relationship between fire fighting abilities and risk preference observed in the case histories given. This was done to investigate whether fictitious information about the relationship between the personality trait such as risk taking and fire fighter ability could produce a perseverant social theory. The case histories reports given to the subjects were manipulated whereby participants were led to perceive that there was either a positive or negative correlation between risk preference and fire fighting abilities. The results demonstrated that participants who were led to believe that risk taking makes better fire-fighters and those initially led to believe that risk taking makes poorer fire-fighters persevered their initial beliefs, even after being debriefed about the fictional nature of the initial information (Anderson et al. 1980). The study showed that the participants adhered to their conclusions even though the evidence fundamental to the conclusions were confounded. Similarly, mock juror studies have found that jurors tend to be unable to disregard evidence that has been ruled inadmissible (Hawkins & Hastie 1990). Equally, in a criminal investigation, the act of considering someone "accountable" (which is a condition necessary for turning a person into a suspect) is in itself likely to increase the belief of the investigator in the culpability of the suspect(s) (Ask & Granhag 2005).

2.1.6. Observer Effect and Contextual Biases

The fact that people can be unaware of pre-existing beliefs has potential consequences in forensic settings. This phenomenon is known as the observer effect, which can be described as when the result of an observation in a particular set of circumstances impacts the observer (W. Thompson 2009a). In forensic science the term observer effect is used when the motives or preconceptions of the observer are thought to influence the perception and interpretation of evidence, resulting in examiner bias (Risinger et al. 2002). Context effect is highly related to observer effect and is used in the forensic sciences to describe situations in which forensic analyses are affected by the context of the crime or by the contextual information available to the analyst prior to their assessment (Saks et al. 2003).

Studies have demonstrated that it is difficult for people to evaluate the strength of evidence independent of pre-existing beliefs and that there is a tendency to devalue disconfirming evidence (Lord et al. 1979). This is because evidence is weighed to support prior beliefs to a greater degree than evidence that contradicts those beliefs (Findley & Scott 2006). The product of various cognitive biases that could obstruct accuracy in what is perceived, how it is perceived, and how it is interpreted is also known in criminal cases as tunnel vision (Burke 2005). Tunnel vision has been shown to have an effect in the initial stages of criminal investigations and this is a significant issue because all subsequent stages of the investigation will potentially be impacted by the information generated at this initial stage (Thompson 2011).

The psychology and social science literature also suggests that people not only demonstrate the outworking of confirmation bias when seeking new information but also in the memory of stored information; meaning that people search their memories in biased ways (Nickerson 1998). For example, in one study, subjects were given a report story to read about a woman whom behaved in number ways that was both extroverted and introverted. After two days, half of the participants were told to assess the suitability of the woman for a job that would reasonably require introversion and the other half was asked to assess the woman's suitability for a job requiring extroverted qualities. The results indicated that those participants asked to assess the suitability of the woman for the introverted job recalled more instances of her introversion. The same effect was demonstrated with the group of participants asked to assess the woman's suitability for the extroverted job where participants remembered more examples of the woman's extroversions (Snyder & Cantor 1979). The participants essentially search their memories in a biased way depending on the suitability of the job required. Studies within memory and recognition show that despite our best intention, memory often fails without our knowledge, with numerous factors affecting how we retrieve from our memory.

2.1.7. Human Memory

The fallibility of memory retrieval have shown that what people tend to remember is likely to be unreliable (Klayman 1995). A study by Roediger & McDermott (1995) showed false memory amongst participants when asked to recall words from a memory game. Participants in this study were presented with a list of twelve words (e.g. awake, bed etc) and were asked straight after to recall as many words as possible. The result showed that on average participants recalled 65 per cent of the presented words. However, most remarkably participants tended to recall and report words that had not been presented with a very high confidence in the 'false memory'. For example the word 'sleep' was reported (although not presented in the list) due to a thematic association with the other words, causing a failure in retrieval of words from a memory task.

The bias in memory and cognition have been shown to occur during the encoding of the memory, the storage of the memory as well as the retrieval stage when the memory is recalled at a later stage (Rholes et al. 1987; Winograd et al. 1998; Green 1992). As mentioned earlier, The world that people know is the one they take in through their minds and senses and normally this does not necessarily resemble an 'accurate' account of the events and the world around us when trying to recall from our memories (Balcetis & Dunning 2006). For example, studies have shown that people have a tendency to overestimate the duration of unpleasant events (Hudson & Nelson 1986). Equally, memory decay over time according to a logarithmic function meaning that events stored in long-term memory are frequently altered and reconstructed based on new beliefs, experiences, and information (Klatzky 1975; Leippe 1980). Furthermore, the way that memories are retrieved can influence their accuracy, for example, the way questions are asked could reduce the accuracy of the recalled memory. In addition, our recollection of feelings of a certain episode will be highly dependent by our current knowledge and feelings about that event (Safer & Keuler 2002; Levine 1997). This could also be affected by incorporating aspects of others accounts of a shared event into our own memories.

Despite deliberate efforts to remember details of an event and having a high level of motivation to remember the event correctly, memories can still be completely undependable. This could have severe consequences in a legal investigation where many convictions subsequently demonstrated to have been erroneous, have been caused by a failure of eyewitness misidentifications (The Innocence Project 2017).

2.1.8. Cognitive Bias and the Legal System

Research regarding cognitive biases and decision-making has also been applied within the legal system. Studies about eyewitness misidentification conducted by Phillips et al. (1999) demonstrated the power of information by indicating that when the suspect is known, it is more likely for the investigator to unconsciously steer the witness towards the suspect. Similarly, research in facial recognition and decision-making has demonstrated that when information is given concerning a suspect with regard to their guilt, people have the tendency to perceive more similarities between a facial composite and the suspect (Charman & Wells 2008).

Numerous studies in policing and interrogations have also shown that people very often fail to attain high levels of performance in making judgments about perception and truth in police interrogations (Soukara et al. 2009; Porter et al. 2000). In fact research has shown that training in the use of verbal and non-verbal behavioural cues for police interrogations has very little or no impact on investigators making accurate judgments of the truth (Kassin & Fong 1999; Frank et al. 2006). Additional studies have also demonstrated variations in interrogation methods when an assumption of guilt had previously been established (Meissner & Kassin 2002), showing that when investigators had a presumption of guilt there was sometimes an unconscious tendency to be more aggressive and intimidating in interrogation towards the suspect, which could potentially increase the risk of false confessions (Perillo & Kassin 2011; Kassin & Kiechel 1996). In fact, police-induced confessions can even be made to appear believable, even when DNA evidence in the case exculpates the accused (Appleby & Kassin 2016). The reports from the Innocence Project show that 1 out of 4 people wrongfully convicted (and later exonerated by DNA) made a false confession or incriminating statement (The Innocence Project 2017). Mock juror studies have also shown that confessions to a crime have more impact on verdicts than other forms of evidence (Scherr et al. 2014; Kassin & Neumann 1997). This is considered to be because most people believe that people do not confess to a crime they did not commit (Kassin 2012). In addition, confession evidence can in fact bias juries, judges as well as forensic examiners (Kassin 2014).

For prosecutors it has also been shown that there are some cognitive pitfalls when involved in an investigation. For example, it has been observed that the prosecution can shape the investigative direction, by determining who to investigate, and once an arrest is made, they determine whether to bring charges or not, what charges to bring and what sentence to seek (Burke 2005) . This approach for prosecutors may lead to potential ways that cognitive bias may impact upon decision-making. Indeed, the phenomenon of confirmation bias could in complex cases lead to the natural tendency to review the case report for confirming evidence and not exculpatory evidence that might contradict the given hypothesis (Findley & Scott 2006). It has also been shown that people can fail to look for evidence that disconfirms a given hypothesis and this can lead to tunnel vision in investigations where investigators could potentially fail to investigate alternative theories of the crime (Thompson 2011). People are motivated to consolidate their beliefs in a manner that strengthens their initial perspective. As mentioned earlier, numerous studies have demonstrated that expectations and motivations can affect how events, people, and evidence are perceived (Bruner & Potter 1964). In criminal investigations this could have severe effects, especially if an individual is being judged by investigators where the initial belief presented to each actor in the system is that the defendant is guilty (Garrett 2008; Garrett et al. 2009).

2.1.9. Miscarriage of Justice

Research and policy makers have started to realise the significant role the science of psychology plays in the study and prevention of wrongful convictions. It is estimated that to date 349 individuals in the United States have been exonerated by post conviction DNA testing, (The Innocence Project 2017). Miscarriages of justice have been identified where there has been a range of causes of error, including fallible eyewitness identification, false confessions, police and prosecutorial misconduct and forensic science error (Kassin et al. 2010; Garrett 2008). Forensic science plays a complex role in the study of wrongful convictions where it has been argued to be both part of the problem but also the solution. For example, DNA evidence has been a major tool to exonerate scores of wrongfully convicted suspects, however in some cases, errors in the DNA evidence were identified. One example of this discussed by Thompson (2010) is the case of Josiah Sutton's (1998) wrongful conviction for rape, where DNA and eyewitness identification was involved in the original case. The analyst testing for DNA in the case was aware that the victim had identified Sutton as one of the rapists. It has been argued that this information may have induced a confirmation bias and led the analyst to focus on evidence supporting Sutton's guilt and ignoring facts inconsistent with that theory. It has been asserted that if forensic scientists are aware of the desired outcome, it is possible that they might unwittingly be influenced to interpret ambiguous data to support a given theory formulated by investigators such as the police and prosecutors (Findley & Scott 2006). The criminal justice system presumes the independence of different types of evidence but these findings suggest that the reality of criminal investigations may not afford such independence, and in some cases the judgments of forensic scientists could significantly be influenced by psychological factors (Kassin et al. 2013).

2.2. The Role of Forensic Science in the Criminal Process

Forensic science concerns the application of science for the purpose of law (Caddy & Cobb 2009). In general terms, forensic science is applied in the investigation of crime, and has become increasingly important in the detection of criminal events, and crime reduction (Mennell & Shaw 2006). The domain of forensic science is varied and draws on a number of different disciplines. Forensic scientists possess knowledge and skills that allow them to collect, analyse and interpret trace materials and evidence associated with, and found at crime scenes. Very often forensic scientists are required by law to communicate their findings to assist courts (Caddy & Cobb 2009), and therefore, maintaining the integrity and security of evidence from its initial discovery to final presentation is crucial (Holobinko 2012).

The conceptual framework outlined by Morgan & Bull (2007) presents six fundamental stages of physical trace evidence within forensic investigations; division and transfer of matter, persistence and tenacity, collection, analysis/identification, interpretation and presentation. Each stage is dependent upon the previous stage being fulfilled, and the framework illustrates the importance of effectively addressing each stage to achieve accurate results and evidence in criminal investigations (Morgan et al. 2009). However, whilst there may be similarities between forensic investigations, the context of an individual crime scene will be specific to that particular event. This context must be incorporated into the appraisal of each crime scene and the complexity of the multiple variables and their relationship to one another acknowledged. This is vital in order to establish the best approach for the collection and analysis of physical evidence and its interpretation in a specific case (Scott et al. 2014).

The value of forensic analysis is well recognised, and the ability of analytical techniques to provide ever more accurate and detailed empirical analysis of forensic samples has been identified (Morgan et al. 2009). However, the interpretation of that evidence in specific forensic contexts is essential. There has been much debate in the literature concerning the methods and approaches that should be taken to offer robust and accurate interpretations of evidence to investigators and to the courts (Fenton et al. 2012). Indeed, a number of cases where the validity of different approaches has been questioned (Redmayne et al. 2011; Fenton et al. 2014) such as the case of R. v

T in the United Kingdom. More recently there has been an augmented awareness of the complexity and uncertainties surrounding the dynamics of evidence that may be recovered from crime scenes. Caution has been called for in the interpretation of physical evidence, with a focus on developing approaches that take into account an empirical evidence base that also incorporates the context specific nature of a particular scene (Morgan & Bull 2007). The necessity for further empirical research within context specific cases has been highlighted where experimental studies, which imitate the forensic reality, are of fundamental importance in order for a measure of the significance of pertinent physical and trace evidence to be identified (Morgan & Bull 2007).

2.2.1. Forensic Science and Expert Evidence

As a result of the complexity of data analysis and interpretation of evidence in the forensic sciences, the issue of admissibility of evidence and expert witness testimonial accounts has been raised (Christensen et al. 2014). Concerns regarding validation and error rates of techniques used by forensic scientists and the professional standards of experts have been articulated in addition to the role of expert witness testimony in court proceedings (Law Comission et al. 2011). In the British and American systems, where trial by jury is the normal state of affairs, the role of the expert witnesses and the evidence that they provide in a courtroom is not only considered as a methodological question, but also an ethical one. It is not the role of a forensic scientist to determine the truthfulness of a variety of propositions related to crime. The role of the forensic scientist is to provide input to the legal process, where the accuracy of the source of various premises pertinent to the evidence presented is made generally by a judge or jury (Thompson 2011). However, it has been documented that experts are often over confident in their abilities, and it has been observed that much of the forensic science evidence presented in court has arguably been accepted without a sufficient degree of scrutiny (Mnookin et al. 2011).

2.2.2. Expert Evidence Standards in the United States

In the majority of American states, the admissibility criteria's applied for expert evidence follows the ruling of the United States Supreme Court in the 1993 case of Daubert v. Merrell Dow Pharmaceuticals, Inc (1993). The admissibility criteria's set out in Daubert are widely known as the Daubert standard (Christensen 2004). Other

American states continue to use the "general acceptance" test established in Frye v. United States (D.C. Cir. 1923). The purpose of the Daubert standard has been to ensure the dependability and significance of scientific or technical expert testimonies admitted in court. The Daubert guidelines allow judges to act as gatekeeper in keeping "junk science" out of the courtroom, and aid judges to evaluate the reliability and relevance of scientific testimonies (Grivas & Komar 2008). The Daubert standard requires evidence presented in court by an expert witness to be testable, subjected to peer review, have established standards, have a known or potential error rate, and be widely accepted by the relevant scientific community (Christensen 2004). The Federal Rule of Evidence (FRE) 702 was appended in 2001 to highlight the connection between the methods and data used, and aimed to focus on the acceptability of the conclusion, rather than the qualification of the expert (Dirkmaat et al. 2008).

The discussion of error and expert evidence intensified in the forensic science community with the publication of the National Acadamy of Science (2009). The report reviewed the standards of process within disciplines undertaking forensic science. The National Academy of Science concluded that there are issues regarding reliability and errors within some forensic disciplines (Found 2014). Furthermore, the report emphasised the potential for subjective interpretation and cognitive bias (Kassin et al. 2013). However, it has been asserted that the forensic community as well as the court often misunderstands the concept of error in this context. Christensen et al. (2014) discuss the difference between scientific error and statistical error rates which have been confused with practitioner errors. Christensen et al. (2014) also highlight the importance for forensic practitioners to ensure that the potential sources or error and limitations within methods used by forensic scientists are not only understood, but also communicated correctly to the legal community.

Furthermore, this was also an issue raised by the PCAST (2016) report which highlighted the importance of acknowledging and addressing all possible sources of error, including cognitive biases, in the forensic science domains (President's Committee of Advisors on Science and Technology 2016). This was recognised as the report emphasised that subjective methods used in forensic science require particularly careful scrutiny due to their heavy reliance on human judgments. This means that the subjective methods are especially vulnerable to human error, inconsistency across examiners, and cognitive biases. The report also concluded that there is a need for clarity about the scientific standards for the validity and reliability for forensic methods. In addition, the report also evaluated specific forensic methods commonly used in the court (DNA, Fingerprint, Bite mark, Hair analysis, Tool mark, Blood pattern analysis), in order to determine if the current methods were scientifically established, to be considered as valid and reliable. The PCAST (2016) report concluded that some of the techniques and procedures used in forensic science lack foundational reliability and consistency. The report also highlighted the need for further empirical studies, as extensive 'experience' in casework cannot substitute for empirical studies of scientific validity.

2.2.3. Expert Evidence Standards in the United Kingdom

In England and Wales, the Law Commission highlighted the issues regarding expert evidence in criminal proceedings in their 2011 report 'Expert Evidence in Criminal Proceedings in England and Wales' where they stated that the judicial approach to the admissibility of expert evidence in England and Wales was passive (Law Comission et al. 2011). In the Criminal Procedure Rules, Rule 33.2 sets out the duty of the expert to the court, with the main objective to provide unbiased objective opinions based within his or her expertise (Ministry of Justice 2013). The Law Commission report however acknowledges that too much expert opinion evidence is admitted without adequate scrutiny, where no test is applied to determine the reliability of the evidence presented. The report proposed that expert evidence in criminal trials should be subjected to a "reliability based admissibility test" before being presented to a jury to exclude unreliable expert evidence. These recommendations were aimed at establishing a framework in criminal proceedings for controlling expert evidence at the admissibility stage, where possible sources of error and bias must be made clear. As a result of the concerns raised by both reports, there has been a call for the development of a research agenda, with the suggested mechanism to enable this being closer collaboration between the professionals within the industry and academic research institutions (Silverman 2011). Whilst this is an admirable aim, there has to date been limited funding made available for primary research within the forensic sciences to address each part of the forensic science process (crime scene investigation, sampling and analysis, interpretation of that analysis and the presentation of evidence in court). However, subsequently, there has been a number of policy reports within the UK highlighting the central role of human cognition in forensic science and its effect on the interpretation of forensic evidence (Government Chief Scientific Advisior 2015; Forensic Science Regulator 2016). These reports highlight the issue of the need to increase objectivity and to minimise cognitive biases entering a criminal investigation at an early stage.

2.2.4. Cognitive Bias and Forensic Experts

The judgments of forensic scientists being influenced by cognitive factors are very different to the effects in investigators' bias, problems in eyewitness identification, and other elements in criminal cases, (as discussed previously section 2.1). The problems in these areas are well known, and jurors (as well as judges) have started to take them into account (Dror 2015). However, scientific evidence presented by experts has a different status. Forensic evidence as well as forensic experts has predominantly been viewed as immune to bias effects, and regarded as objective and impartial (Mnookin et al. 2011).

Experts often have specific cognitive abilities needed to perform certain tasks associated within their expert domains (Dror 2016). The 'expertise' of an expert has been acquired by repeated exposure to the tasks they need to perform, creating schemas from learning and experiences (Dror 2011). Indeed, experts reliance on top-down information (as discussed throughout Chapter 2.1) allows for enhanced, quicker and efficient performance, learning how to 'automatically' filter out noise and dealing with large amount of information (Edmond et al. 2016; Stanovich 2014). This leads to experts being able to perform skills effortlessly. Paradoxically, the cognitive architecture involved in being an expert could also result in lack of flexibility resulting in experts missing or ignoring important information, resulting in overconfidence in their abilities (Sternberg 2002). This could be problematic specifically in domains including uncertainties and risk taking (Dror 2011). Meyer & Booker (1991) and Dror et al. (2005) highlighted the mental cognitive process behind the opinion of an expert known as elicitation. This consists of four cognitive tasks:

- defining the question,

- remembering the accurate information,
- making a decision and
- reaching a conclusion.

In a forensic context this would be known as "what is classified as evidence, what is recognised as collected evidence, and what is examined and how it is interpreted." (Cooley & Turvey 2011 p.69). The expert must first understand what has been asked of them in order to answer a question. This demands a specific focus on the accurate information, and the limitation of personal speculation (Meyer & Booker 1991). However, when an expert tries to consider the accuracy of information, different cognitive factors (as mentioned previously) will play a vital role, and might cause selective attention towards information causing an observer effect (Budowle et al. 2009). This essentially means that what is remembered and perceived by the expert depends upon the perceiver themselves (Blackwell & Holmes 2010). The power of schemata and other combined cognitive processes will affect each expert individually for what is remembered as accurate data. Forensic context generally involves large amounts of multivariate information, often too complex for one individual to process (Fraser-Mackenzie et al. 2013). The decision-making of an expert is also dependent upon the manner in which problems are structured and presented. The same problem can result in different decisions depending on how the problem is framed and displayed (Phillips et al. 1999). For example, studies have shown that forensic experts will evaluate evidence differently depending upon whether they are consulting for the prosecution or defence (Murrie et al. 2013). Internal and external factors (as discussed in previous sections) could affect the decision-making outcome. It has also been demonstrated that the internal factors will vary at different times, which can cause the same expert to change their judgment on the same identical decision (Dror & Charlton 2006).

Furthermore, as mentioned in section 2.1.7, despite best intentions, memories are fallible. Compared to novices, experts rapidly retrieve from memory, previous instances and decisions relevant to the current situation (Ericsson et al. 1994). Although this can be useful, the specialised nature of expertise can also render experts inflexibility and be especially prone to external influences as well as searching memory in a biased way (Edmond et al. 2016). In forensic science this means that for example "having performed ten-thousands autopsies might not enable a forensic pathologist to accurately recall the frequency of a particular type of stab wound" (Edmond et al. 2016 p.5). Much of the activities that forensic experts (as well as other experts) perform on an everyday basis are done automatically and with little cognitive effort (Dror et al. 2012). This means that there is very limited access to the cognitive processes that determine the choices being made. This is also known within psychology as the introspection illusion (Nisbett & Wilson 1977; Wilson & Dunn 2004; Johansson et al. 2006) with dozens of experiments on self-assessments showing a weak relationship between actual performance and self-rated performance (Anders Ericsson et al. 1993; Ericsson et al. 1994). This could significantly have an effect on expert testimonies, as expertise in a specific domain does not necessarily include the ability to articulate the basis of the reason for a decision or action. Communicating expert evidence is considered to be highly difficult as well as error prone as what experts might say may not be what a lay audience hears (Edmond et al. 2016).

Evidence of miscommunication has also been found in cases where numerical rather than verbal expressions are relied upon. The end result for a forensic scientist is to reach a scientific conclusion based on the relevant evidence and to communicate the results in a manner that can be understood by non-scientists (Springer 2007). There is a growing acceptance by the forensic science community of the value of probability frameworks as a means to offer a comprehensible format for the formulation and presentation of opinions in forensic science, such as the estimation of a likelihood ratio (Martire et al. 2013). Uptakes of these standards have varied considerably across disciplines and jurisdictions. For example in the Netherlands, the likelihood ratio is considered as standard practice for bullet comparison and is actively being expanded to other disciplines. However in the United States, likelihood ratio comparison is not common in disciplines other than DNA analysis (Kassin et al. 2013). The debate about the best way to present the results of complex forensic analysis in court has not included the body of empirical evidence collected by psychologists in the decision-making field and reasoning under uncertainty. Numerous psychological studies suggest that people often have difficulties understanding probabilistic and statistical estimations (e.g. Gigerenzer & Gaissmaier 2011).

Furthermore, the degree and content of the contextual information forensic scientists should know about a case has only started to receive attention in the forensic science literature (Mattijssen et al. 2016; van den Eeden et al. 2016; Dror 2014a). Some commentators have argued that ignorance of the facts of a case may cause forensic scientists to ask and answer the wrong questions, which could potentially be harmful to an investigation (Champod 2014). However, cross communication could potentially affect all stages of the elicitation task involved in a forensic investigation and cause judgment and decision-making to be affected at each stage of the forensic process from the initial analysis to the court (Thompson 2011). The four cognitive elicitation tasks (defining the question, remembering the accurate information, making a decision and reaching a conclusion) are all relevant to the forensic conceptual framework (Morgan & Bull 2007; Inman & Rudin 2002), and every expert in the field of forensic science. As a result there has been a rise in interest across the forensic science domains as to which stages cognitive biases may arise during an investigation, with empirical research being conducted within different forensic domains to investigate these effects (Found 2014).

2.2.5. Cognitive Bias and Research in Forensic Science

Studies conducted to assess the cognitive processes and the tendency for biases within human decision-making in the forensic field are being undertaken within a number of forensic domains. Research has indicated that human error due to cognitive patterns can influence and cause a reduction in the objectivity of forensic experts when analysing evidence (Thompson & Cole 2007; Kassin et al. 2013; Fraser-Mackenzie et al. 2013; Dror 2011). Various factors such as extraneous context, time pressure expectation, and motivational statements have been shown to have an influence on observation and decision-making (Dror & Fraser-Mackenzie 2008).

For example, social interaction, past experiences and prior information has been argued to influence forensic handwriting and document examinations in their final conclusions (Miller 1984; Stoel et al. 2014a; Kukucka & Kassin 2014). The result of the research conducted in this area has demonstrated that the preconceived notion that ‘the suspect wrote or did not write the document’ affects judgments and decision-making and ultimately the final conclusion of the forensic document examiner (Miller 1984). In addition, the effect of contextual information has also been shown within fingerprint examiners with regards to whether or not two fingerprint marks originate from the same source (Dror and Charlton, 2006; Dror, Charlton and Péron, 2006; Dror et al. 2011). In many of these experiments, the majority of experts reached different conclusions on previously assessed finger mark comparison, and were inconsistent in their analysis when provided with new contextual information and whilst undertaking new visual imaging (Dror et al. 2005; Dror and Charlton, 2006), showing potential for confirmation and contextual biases. Findings from other studies within fingerprint comparison and contextual biases also found that even without the context of the comparison print there was still a lack of consistency in analysing some latent marks (Schiffer and Champod, 2007; Langenburg et al., 2009; Dror et al., 2011). Not only was this reflected by inconsistency between different experts, but also the same experts at different times were inconsistent with their own analysis (Dror et al. 2011). This shows a lack of reproducibility between and within examiners even when context is absent. In addition, studies has also shown that the position of matching prints in the 'line up' affected the way in which fingerprint examiners assessed the print, highlighting the degree of false exclusion and inconclusive identifications across a series of mark evaluations (Dror et al. 2012)

Similarly, contextual effects have also been empirically studied within the domain of DNA where the interpretation of a mixed DNA sample differed amongst DNA experts depending on the case context (Dror and Hampikian 2011). The result from this study showed that when the DNA mixture (taken from an adjudicated criminal case involving a gang rape) was presented to 17 neutral DNA examiners (with no contextual information or case background provided), only 1 expert agreed with the original examination. Four of the DNA experts stated the sample to be inconclusive and 12 excluded the suspect in question. Equally, the effect of context and potential for confirmation bias has also been identified within forensic anthropology (Nakhaeizadeh, Dror, et al. 2014; Nakhaeizadeh, Hanson, et al. 2014; Klales & Lesciotto 2016) Bite mark comparisons (Page et al., 2011; Osborne et al., 2014), bloodstain analysis (Osborne et al. 2016), forensic entomology (Archer & Wallman 2016) and fire scene examinations (Bieber 2012).

Studies have also been conducted within motivational and emotional factors that might influence the performance of a forensic expert (Charlton et al. 2010; Dror et al. 2005). The results from these studies indicated that although confirmation bias normally operates outside of conscious awareness, forensic examiners might have some insight into the cognitive motivational and emotional factors that may affect the decision-making process. For example, Charlton et al. (2010) conducted a series of semi-structured interviews of fingerprint examiners where the examiners expressed a personal interest in solving crime and catching the offenders, showing a cognitive motivational factor when assessing forensic fingerprints.

Training and experience could also have an effect upon expert decision-making, and the individual differences will characterise the degree to which a particular context will affect an expert. Research has demonstrated that diverse manipulations of context may affect people differently and it is often in ambiguous cases where the levels of cognitive bias will have the most impact on the outcome (Thompson & Ford 1991) . Kerstholt et al. (2010) presented a study on bullet analysis with the intention to observe whether additional incriminating contextual information would affect the expert when observing similarities between two bullets. The results, however, indicated that the contextual information given in the case had no effect on the conclusion. It is therefore important to acknowledge that contextual bias may affect the process but not necessarily the decision-making outcome of the forensic examiner (Kerstholt et al. 2010).

2.2.6. Cognitive Bias and Casework

The issue of bias and cognitive vulnerability has also been demonstrated to be some of the sources of error in high profile forensic cases such as those of Shirley McKie in Scotland (Dror & Cole 2010), Brandon Mayfield in the US, (Stacy 2006) and Amanda Knox in Italy (Kassin 2012). Commentators on these cases have asserted that it is important to acknowledge the numerous pitfalls that can occur within decision-making when justice agencies arguably work too close together, and where tunnel vision, social conformity, group thinking and context biases can have significant influences resulting in a chain of biased interpretations (Dror & Cole 2010; Kassin et al. 2013) In the Shirley McKie case, the Scottish government specifically set up the Fingerprint Inquiry (under the Inquiries Act 2005) to address the steps, taken to identify the fingerprint associated with McKie, which led to the case of HM Advocate v. McKie 1999 (Fingerprint Inquiry 2011). The goal of the inquiry was to report on findings of fact and determine the consequences of steps taken in this case as well as provide recommendations for the future. The inquiry report was published in December, 2011 with one of the findings articulated in the report concerning the decision-making processes in fingerprint analysis and the manner of presenting analysis conclusions.

2.2.7. Addressing Cognitive Bias in Forensic Science and Criminal Investigations

Even though there is a growing acceptance of the role of cognitive biases and its implications in forensic science and criminal investigations, in practice, procedural changes to counter cognitive issues do not seem to have been structurally implemented (Stoel et al. 2014). One of the potential reasons for this could be the misinterpretation of cognitive biases as being an ethical issue. Cognitive biases occur without awareness or intention and are the predictable result of the human cognitive and psychological systems, rather than intentional misconduct. It has been demonstrated that cognitive biases cannot be conquered by will- power, as it is not possible to be fully appreciative of the extent to which people are affected by cognitive errors (Thompson 2009). Although education in human cognition could potentially improve the decision-making of an expert, it is not possible for education alone to minimise and reduce cognitive biasing effects (Kassin et al. 2013). Therefore, a number of different approaches have been identified as means of addressing cognitive bias in the forensic sciences and the criminal investigation system.

2.2.8. The Legal System

Within the legal system one of the proposed solutions for prosecutors is to incorporate the practice of providing pro-defence counterarguments to the prosecutorial interpretation of the evidence against the defendant (Findley & Scott 2006). Generating explanatory counterarguments can mitigate belief perseverance by simply switching between prosecution and defence mind-sets to produce plausible explanations of both guilt and innocence for each piece of evidence (Burke 2005). Other solutions within law enforcement (other than educating judges, prosecutors and defence lawyers about cognitive biases) have been to include additional unbiased decision-makers in the process by providing 'fresh look reviews.' (Burke 2005). In addition, solutions have also been proposed regarding legal decision makers being educated within the procedure by which the forensic examiner reaches their conclusion (Kassin et al. 2013). This is an important step given that the decision-making processes during evidence collection, analysis, and interpretation are likely to be strongly related to how evidence is presented and evaluated in court.

Cross talk and information exchange between different units of the justice system occurs routinely in forensic investigations. However, too much communication of irrelevant information at the earliest stages of a crime scene investigation has been argued to potentially lead to system failure (Thompson 2009; Saks et al. 2003). The National Academy of Sciences in the United States has reported that crime laboratories should not fall under the umbrella of law enforcement, which is the case in some other countries and jurisdictions (Stoel et al. 2014a). For example, Washington, D.C. formally separated its laboratories from the police and instead established the District of Columbia Consolidated Forensic Laboratories. The consequence of law enforcement agencies collaborating too closely with each other creates the risk of cognitive biases altering the judgment and interpretations of an expert at the initial stage of a forensic investigation. For example, if analysts are exposed to contextual facts regarding the crime there is the potential for the effective 'double counting' of evidence. This may occur if the analyst is influenced by the evidence of a confession in the determination of uncertainty regarding a possible match of a fingerprint which could lead the jury to think they are receiving two independent pieces of evidence (confession and fingerprint evidence), as they are unlikely to know that the result of the print analysis was affected by the evidence of a confession (Thompson 2011).

2.2.9. Case Manager Model

Some of the proposed solutions to minimise cognitive influences and prevent double counting of evidence in forensic science have been to separate various laboratory functions by assigning them to different people (Thompson 2011). One suggestion is to apply a case manager model. The role of the case manager typically includes communications with police officers, participation in the decisions of what specimens to collect at a crime scene and what tests to run. Case managers will therefore be responsible for placing the test results in context and assessing the importance of forensic observations with various theories of what occurred (Dror 2014). Such an approach allows case managers to understand the context of a case and analysts to be blind to domain irrelevant context and thereby protected from contextual bias. Similar solutions have been proposed by Saks et al. (2003), who proposed the creation of evidence and quality control officers (EQC), who could act as highly trained individuals within exhibit management units. Their main responsibilities would be to filter out domain irrelevant information, formulate the questions to be answered in the least suggestive way, and coordinate the submission of the evidence to the appropriate section (Saks et al. 2003). It is crucial for the 'success' of any forensic analysis, interpretation and presentation that the collection of evidence is carried out accurately and appropriately (Morgan & Bull 2007). By adopting these models, crime scene collections, sampling procedures, and analysis have the potential to be shielded from cognitive factors to a greater degree. This will strategically separate (to the best of our abilities) judgments and evaluations from being contaminated by cognitive biases at the earliest stage of an investigation. In addition, it will also allow forensic scientists to extract contextual knowledge that is of relevance. It is asserted that a blind procedure will only eliminate domain irrelevant information, allowing forensic scientists to deal in an effective way with the complexity and uncertainties involved at a crime scene (Mattijssen et al. 2016; Dror 2014; Dror et al. 2015)

2.2.10. Laboratory

It is understood that in forensic laboratories, the decisions, interpretation and verification stages could also be affected by human factors. In DNA analysis, sequential unmasking has been suggested as a hybrid approach to minimise the potential for contextual bias, where a known DNA profile might affect the interpretation of an complex and mixed DNA evidence sample (Krane et al. 2008; Thompson 2009; Dror et al. 2015). It has been suggested that this approach addresses the issue by offering the means of analysts making an initial examination of samples prior to learning the profiles of suspects or known contributors (Dror et al. 2015). However, the verification stage also needs to be considered when combating cognitive biases. In many forensic laboratories verification stages are mainly performed on positive identifications, potentially causing base rate regularities (Dror 2014) Very often the second examiner verifies the first examiners work knowing the decision-making outcome. One proposed solution includes blind verifications, whereby the verifier does not know the conclusion of the first examiner, and is unaware of what decisions they are verifying (Dror et al. 2015). Another potential solution suggested to enhance accurate judgments and decision- making in forensic science techniques using match judgments (Dror & Cole 2010), (such as DNA analysis and fingerprint examination) is the filler control method . This approach provides forensic examiners with a minimum of three samples rather than two for comparison, including a crime scene sample, suspect sample and filler(s) samples. It is suggested that this method will enable the forensic examiner to know which sample is from the suspect and which are from the fillers (Wells et al. 2013), thereby protecting examiners from contextual influences in the estimation of error rates for the techniques used as well as the individual analysis. It is also important to acknowledge that not all laboratories have the resources or time to apply all these procedures (Langenburg, 2017). Therefore, solutions have been proposed in the form of adopting a triage approach where each laboratory assesses the case in question and assigns resources where they are needed (Dror 2014). The degree of vulnerability to cognitive bias is dependent upon the complexity of the case (i.e., how difficult it is, how near it is to the decision threshold) as well as to the level of exposure to biasing information; each laboratory can use the triage approach to classify cases into different procedures (such as the level of blind verification) according to their vulnerabilities to bias.

Discovering the different predictors of bias causing interpretation issues within each forensic domain is also an important factor. For example in the fingerprint domain, quantitative image measures for estimating error rates have been applied to discover objective predictors of error (Kellman et al. 2014). Within the fingerprint domain estimating an overall error rate can be challenging, though some fingerprint comparisons may be more accurately compared to others that are historically more prone to bias interpretations. The study by Kellman et al. (2014) indicated that the distribution of error rates varies depending on the visual content of the specific comparison. It highlighted how the difficulties of assessing fingerprints might impact on how judges and juries understand the admissibility of a specific fingerprint comparison. The study also outlined the underlying factors that make some fingerprints more difficult to compare to have a strong impact upon the training of fingerprint experts and the selection of examiners. The study advocates that forensic examiners need to have the cognitive ability to perform the task given to them and that developing tests that specifically focus and quantify these abilities are needed in any forensic domain in order to better allocate resources (Dror 2014) .

Technological solutions to address cognitive biases could potentially be very useful. A good number of recent studies in forensic science are now based on new metric methods where statistics, algorithms and technology are applied. The increase of forensic technology has greatly improved forensic work. However, it is important to acknowledge the new spectrum of cognitive challenges these technologies might provide. For example, as mentioned earlier, the use of the AFIS system could potentially create base rate regularities amongst the expectations of the experts (Dror et al. 2012). Huge searches on databases could also create a higher chance of finding incidental similarities when comparing if a mark from a crime scene comes from the same source as known marks (Mnookin et al. 2011).

Implementation of valid solutions in the combat of cognitive biases in the forensic domains currently varies considerable between laboratories, countries, and domains. For example, laboratories such as the FBI and NIST have modified their standards and procedures to minimise biasing effects (President's Committee of Advisors on Science and Technology 2016). This has also been followed by the National Forensic Institute in the Netherlands in domains such as fire arms investigations (Mattijssen et al. 2016). Although many forensic domains are considering the proposed solutions for minimizing cognitive biases (Krane et al. 2008; Dror et al. 2015; Archer and Wallman, 2016; Langenburg, 2017) in some domains such as forensic anthropology very little empirical studies has been undertaken to establish the extent of cognitive issues specifically pertinent to forensic anthropological approaches.

2.3. Forensic Anthropology

The forensic anthropology domain brings together the techniques, methods, and application of physical anthropology to medico-legal questions (Dirkmaat & Cabo 2012). In recent years, forensic anthropology has experienced significant expansion and growth, with increased professional interest and public attention (Christensen & Crowder 2009). Indeed, the discipline has become increasingly recognised by law enforcement and investigators as having a valuable role to play in the generation of intelligence and evidence. However, variation and development in forensic anthropology as well as the remit of the forensic anthropologist differs significantly between countries and jurisdictions (Kranioti & Paine 2011; Ubelaker 2015). In most countries forensic anthropologists particularly work with skeletal human remains, in addition to analysing bone from fleshed, decomposed or burnt remains. Whilst in other countries some forensic anthropologists also work with living populations (Cattaneo 2007). This means the role of the forensic anthropologist varies significantly depending on country, education and training.

Generally speaking, forensic anthropologists are trained to provide an osteobiography, (also known as a biological profile), by applying methods that will assist in the assessment of the sex, ancestry, age at death and stature of an individual. In addition, forensic anthropologists contribute to the recovery of human remains and to the classification of individual variation, which assist in determining the identification of unknown individuals, and in some cases the cause and manner of death (Pickering & Bachman 2009). In addition, forensic anthropologists might also be asked to comment on the post-mortem interval, time since death, or any taphonomical processes in order to reconstruct the events surrounding a death (Márquez-Grant 2015). In some countries archaeological methods applied to forensic and medical legal cases usually fall under the practice of forensic anthropology (Márquez-Grant 2015). However, in other countries, the discipline of archaeology and anthropology are considered to be separate, with each discipline having their own accreditation system (Groen et al. 2015).

Recently, there has been an increase in the critique of some of the techniques used by forensic anthropologists. Discussion has been extensive concerning evidence validation, admissibility, and error rates in the methods applied (Christensen &

Crowder 2009; Cattaneo 2007). Whilst many of the issues that have been identified have been addressed, with some modifications to existing qualitative methodological approaches, (Mahakkanukrauh et al. 2016; Clark et al. 2015; Hefner et al. 2015; Lottering et al. 2013) some areas are yet to be fully addressed. The presence of cognitive bias, its impact and the cognitive processes involved in the assessment of human remains have only recently begun to be assessed in the published literature (Nakhaeizadeh, Dror et al. 2014; Klales & Lesciotto 2016; Nakhaeizadeh, Hanson et al. 2014).

2.3.1. The Biological Profile

The general practice of forensic anthropology addresses the establishment of a biological profile of human remains, which traditionally consists of sex, ancestry, age at death, stature, pathology and trauma assessments (Dirkmaat & Cabo 2012). The method used to evaluate each parameter is achieved by the application of metric and nonmetric techniques. Metric assessments are based on measurements of skeletal elements, and non-metric assessments on qualitative observation techniques of gross morphology of skeletal characteristics (Pickering & Bachman 2009; Cattaneo 2007; Cunha & Cattaneo 2006; Christensen & Crowder 2009).

It has been argued that non-metric techniques used in forensic anthropology are generally reliant upon observation and the specialised experience of the observer (Hefner et al. 2007; White & Folkens 2005; Dirkmaat et al. 2008). Some have even contested the techniques, and asserted that they are limited because of their subjective nature (Walrath et al. 2004; Lottering et al. 2013; Spradley & Jantz 2011) in a manner akin to other forensic disciplines. In response, there has been some modifications of existing methods that have accompanied the development of new comparative samples in forensic anthropology and statistical tools for data analysis (Spradley & Jantz 2011; Walker 2008; Hefner & Ousley 2014; Dirkmaat & Cabo 2012; Grivas & Komar 2008). These developments have enhanced the role of quantitative methods and have led to a rise in new publications in the literature concerning the analysis of skeletal remains (Klales et al. 2012) . Furthermore, attempts to standardise the diversity of methods used in biological profiling and data collection have started to be considered, in particular with the work of expert groups (e.g., the Scientific Working Group of Forensic Anthropology (SWGANTH), now reorganised under the National Institute of Standards and Technology (NIST) as Scientific Area Committees (SAC)) (SWGANTH 2016).

There is still considered to be variation and inconsistency among practitioners in the way in which methods are employed and how the results are reported when establishing a biological profile (Bruzek 2002). Most quantitative techniques have still not been widely adopted across the discipline with anthropologists still preferring to use the traditional non-metric observation of morphological traits in sex, ancestry, and age at death estimations, all of which could be susceptible to cognitive interpretation issues (Nakhaeizadeh & Morgan 2015).

2.3.2. Human/ Non-Human

One of the first steps in the identification process for a forensic anthropologist when observing skeletal material of unknown origin is to determine if the bone in question is human or non-human (Sorg & Haglund 2001). In some cases human skeletal remains are presented as highly fragmented, damaged and potentially commingled with non-human skeletal remains (Cattaneo et al. 1999). Therefore, identifying fragmented or isolated bones in a forensic context can be challenging, and some have argued that observation of gross morphology (the most common method used to distinguish between human from non-human) is highly related to the experience of the anthropologist (Dominguez & Crowder 2012). Human remains can often be mixed with those of animals and it has been acknowledged that the identification can be further complicated by modifying factors (Hillier & Bell 2007). In some cases the forensic anthropologist may also use histological approaches as well as DNA in order to distinguish between human and non-human remains. However, DNA is considered too time consuming, destructive and not cost efficient (Dominguez & Crowder 2012). Therefore, the majority of interpretation of distinguishing human bones from non-human is based upon visual examinations, arguably being considered as highly subjective.

2.3.3. Sex assessment

The first step when generating a biological profile of an unidentified individual is the estimation of sex (Guyomarc'h et al. 2011). This is primarily owing to some of the traditional methods applied for age estimation, ancestry, and stature being sex specific (Klales 2013). For example, the observable differences in aging and growth patterns between sexes, as well as variations in morphological traits related to ancestry, makes accuracy of sex estimations vital (Krishan et al. 2016). Many have argued that the accuracy of sexing of skeletal remains depends much on the element present for analysis and its preservation state (Naikmasur et al. 2010; Đurić et al. 2005) The most extensively adopted sexing techniques are based on morphological observations, and rely on the visual assessment of sexually dimorphic traits (Mahfouz et al. 2007). These assessments are generally used by forensic anthropologists, due to their efficiency as well as their practicality (Biwasaka et al. 2012; Đurić et al. 2005). However, the methods used in sexually dimorphic traits have been argued to be influenced by the level of subjectivity (Steyn et al. 2004; Kemkes-Grottenthaler et al. 2002). In addition, visual assessments in sex estimations generally show higher accuracy results with intact bones as the level of accuracy tends to decrease in incomplete and fragmented skeletons (Krishan et al. 2016). Nevertheless, some key features of sex assessment (e.g. the shape of the pubic bone) can only be assessed morphologically (Patriquin et al. 2003).

Non-metric assessments in sex estimations generally involve visual evaluations of traits against provided descriptions and illustrations. Historically, the bones of the pelvis and the skull have been the most frequently used elements for sex estimation, with the bones of the pelvic area considered as the single best indicator of sex (Spradley & Jantz 2011). For example, one of the sexually dimorphic characteristics in the pelvic region that is considered to differ between males and females is the sub-pubic angle, which is considered to be larger for females owing to childbearing. Most methods conducted in sex estimation on the pelvic are focused on the part of what is known as the innominate bone (hip bone), which is also referred to as the os coxa. Much of the current studies on the innominate bones are expanding upon the morphological traits outlined by Phenice (1969). Visual assessments on the pelvis have been shown to provide a variation in classification accuracy depending on what area and traits within the pelvis are being studied. The majority of the studies have shown a classification accuracy above 85% (Sutherland & Suchey 1991; Klales et al. 2012; Rogers & Saunders 1994; Bruzek 2002; Yaşar Işcan 1988) with variation in intra-observer errors.

Non-metric traits of the skull are also being used for sex determination on skeletal remains. Unlike the pelvis, (which has sexually dimorphic features mainly due to reproductive differences), the skull has shape and trait differences that are mainly due to size (Byers 2010). For example, very generally speaking, the female crania is considered to be smaller and more gracile relative to a male (Garvin 2012). Similar to the methods developed for the pelvis, sex assessments on the skull are based on observations of multiple traits. Generally, visual scoring features of the skull are conducted by using a rating scale (normally between 1-5) for several important features (Bukistra & Ubelaker 1994). The method allows for a visual observation of each trait with the scoring of 1 representative of a female (minimal trait expressions) and 5 being male (maximum trait expressions). The accuracy of the visual methods applied to the skull has reported varying accuracy rates with some studies showing correct sex classification for some cranial traits to be above 80% (Garvin 2012). However, visual scoring features of the skull has also shown a degree of subjectivity with divers inter-observer scores reported within the literature (Lewis & Garvin 2016; Walrath et al. 2004; Walker 2008; Williams & Rogers 2006). In addition, morphological traits on the skull have been argued to differ amongst and between populations with intermediate forms of trait expressions being observed (Garvin 2012). Recommendations have been made for knowing the population being studied when applying the methods for sex assessments on the skull (Lewis & Garvin 2016). Others have also argued that most postcranial elements outperform the skull in estimating sex, showing a much higher accuracy rate (above 90%) when combining metric methods with multivariate discriminant function models on postcranial elements (Spradley & Jantz 2011).

Metric sex assessments are based on the basic principle of variability between male and female dimensions and have been acclaimed for their objectivity and their much lower intra- and inter-observer variation of individual interpretation (Soni et al. 2010; Pretorius et al. 2006; Asala et al. 2001). Metric assessments have been acknowledged to enable the easier application of quantitative statistical analyses with associated error and probabilistic estimations (Giles & Elliot 1963; Kimmerle et al. 2008; Spradley & Jantz 2011). Accuracy in traditional osteometric methods for sex estimations has been shown to range between 85-95 % (Krishan et al. 2016; Scheuer 2002). However, the accuracy in sexing based on metric assessments may vary significantly depending on the statistical model utilized (Krishan et al. 2016). The most common statistical approach used in osteometric analysis for sex estimation in unidentified skeletal remains is the use of Discrimination Function Analysis (DFA). The use of DFA has become a popular trend in the determination of sex of unidentified remains, and is considered to be a relatively simple method to use (Franklin et al. 2008). The results, however, of DFA depend on the sample size and sexual dimorphism of the population for which the DFA have been developed (Bidmos et al. 2010).

Further, it is difficult to attribute differences in size to sex without considering ecological and physiological implications (Garvin 2012). Not only are metric analyses limited owing to issues of the variation in size within pertinent populations, but this form of assessment also requires intact skeletal elements. This is not always possible in forensic cases where skeletal elements can be incomplete or damaged due to various reasons (Biwasaka et al. 2012). It is worth noting, however, that while traditional metric methods are objective in essence, they often also suffer from observer discrepancies if landmarks are not properly defined (Krishan et al. 2016).

In spite of continual improvements in sexing methods in both existing as well as introduction to new approaches (Kranioti et al. 2009; Mahakkanukrauh et al. 2016; Biwasaka et al. 2009), there is still a need for generating population specific standards with reliable population based data. Although improvements have been made within the discipline to produce more objective methodological approaches (specifically within metric assessments) in sex estimations, traditional non-metric methods are still widely cited and used by forensic anthropologists.

2.3.4. Ancestry assessments

The use of diverse morphological skeletal traits for the estimation of ancestry is an important part of the forensic anthropologist's establishment of a biological profile. In physical anthropology, non-metric analysis has been utilized in an effort to categorise human groups, with traditional approaches to ancestry assessment relying predominantly on observations of presence or absence of morphological traits (Hefner 2009). Non-metric ancestry assessments include observing and scoring the morphology of the skull, (including the facial bones, and the mandible), which are visual identifications of traits that are thought to differ between groups (Byers 2010) and are mainly based on trait lists (Walrath et al. 2004). The most common traits described and used in textbooks and training manuals in forensic anthropology have derived from the trait list outlined by Rhine (1990). Generally speaking, the trait lists developed (following Rhine (1990)), categorise human groups in three discrete racial groups, White, Asian, Black, with a more recently added fourth group, Hispanic (Spradley et al. 2008).

However, non-metric ancestry analyses have not been scrutinised to the same level of standardisation as metric analyses but are, nonetheless, more widely used by forensic anthropologists (Hefner et al. 2012). It has been debated that the traditional non-metric assessment of ancestry is highly subjective, and has been determined not to be reliable when conducted by visual observations alone (Rubin & DeLeon 2017; Hefner 2009). In addition, it has also been argued that the non-metric methods are predominantly based on observer experience rather than the consideration of distribution of morphological traits among modern populations (Hefner et al. 2007). It has been claimed that relying on typological, experience-based approaches has produced a method that is "as much an art as it is science" (Hefner 2009 p.1). In addition, comparisons of ancestry based on typological trait lists have been criticised, as they are established on extreme trait expression and trait distributions. Some have argued that these trait distributions and expression have not yet been empirically supported (Hefner et al. 2012). Much of the critique has been directed to the small sample size used, as well as the distribution of the sample size for each ancestral group. For example, only 87 skulls were included in the sample size used in Rhines 1990 study with only seven skulls examined for the 'Black' ancestral group (Hefner et al. 2012). Nonetheless, many forensic anthropological training manuals and taught programmes still prefer to cite and use the traits outlined by Rhine (1990) (Burns 1999; Byers 2010).

Furthermore, issues regarding human judgment processing, which could greatly diminish the value of non-metric trait analysis in ancestry assessments, have also been highlighted as a concern within the literature (Hefner et al. 2012). Therefore, the potential effects of contextual influences may be considerable. However much improvement is being made within non-metric assessments in ancestry in the forensic anthropological domain with new comparative samples and studies within morphoscopic trait expressions (Hefner et al. 2015). For example, suggestions of using classification statistics of non-metric traits in ancestry assessments have demonstrated a reduction in the human judgment process and a classification rate above 83% (Hefner 2009).

Metric methods and standard cranial measurements are (to some extent) commonly used for ancestry estimations. Much development has taken place in the collection of population based cranial landmarks and data collection for ancestral assessments (Spradley & Jantz 2016). The most common software programs in use for ancestry estimations based on metric assessments are Fordisc 3.0 and CranID.

Fordisc 3.0 is a computer program that employs multivariate statistical classification methods to estimate sex, ancestry and stature of unknown skeletal samples, using various anthropometric measurements (Dirkmaat et al. 2008). The data behind the software largely originated from the Forensic Data Bank in order to record information about modern populations, primarily from forensic cases (Krishan et al. 2016). Although widely used by forensic anthropologists, this tool has been recently criticised, primarily when used to determine ancestry (Guyomarc'h et al. 2011). Some of the critique has included the lack of individual ethnic groups being presented, as the program tends to be more useful when applied to skeletal remains belonging to one of the populations presented in the reference sample (Elliott & Collard 2009; Urbanová et al. 2014). Furthermore, the mixed population and the continued 'genetic exchange' between different 'ethnic' groups can cause miscalculations owing to gene overlap (Krishan et al. 2016).

CranID was developed to determine geographical origin of skeletal remains in archaeological and forensic contexts. Similar to Fordisc, CranID uses multivariation linear discrimination analysis to assess the geographical origin of unknown skeletal remains, using measurements from the cranium. Some criticism that has followed is the inter and intra observer error with the measurements taken (Smith 2012). Similar to Fordisc, much of the criticism of CranID remains within lack of individual ethnic groups presented in the reference material.

2.3.5. Age at Death

Age at death is another important parameter in the forensic identification process. The estimation of age at death of adult skeletons continues to be a developing area in the field. Age-related bone change is highly dependent on variability between individuals and other complex factors, resulting in different changes in the bones depending on the population being studied (Martins et al. 2012; Hens & Belcastro 2012).The majority of methods applied in age at death estimations are based on visual scoring of morphological indicators. These indicators include changes in the bone surface as well as different fusion stages of the bones, with a common approach being to assess the changes in the different parts of the pelvis bone and sternal extremities of the ribs (Lottering et al. 2013; Chen et al. 2011).

Recently, there has been an increase in the revision of current methods applied in age at death assessments–(Hartnett 2010). Some have highlighted the need to develop new methods and review existing ones with contemporary reference samples (Lottering et al. 2013; Langley-Shirley & Jantz 2010; Márquez-Grant 2015). This has increased the interest in expanding current reference collections in age at death, with new reference samples increasing not only in North America, but also in Europe, Australia, Asia, and Africa (Márquez-Grant 2015). Efforts have also been made to collect new population based comparative landmarks for age at death (see Passalacqua (2009) for the *sacrum*, Rissech et al. (2007) for the *acetabulum,* and Falys & Prangle (2015) for age changes in the *clavicle*). Furthermore, a number of histological approaches have also been undertaken in the development of age at death assessments (Crowder & Stout 2011).

Despite the increase of research in new methodologies, the most commonly used method in age at death assessments in adult remains is the changes observed on the pubic symphysis, originally developed by Brooks and Suchey (1990). The theory of this method relies on the concept that over a lifetime, the surface of the pubis changes in predictable age related ways. The method relies on visual assessment of the different stages related to age growth. The method have been considered to be one of the most reliable age at death methods to use, especially for those under the age of 40 (Márquez-Grant 2015). The Brooks and Suchey (1990) method can be employed in a number of ways. The different age stages of the pubic symphysis can be observed by using the written descriptions followed for each age stage, series of casts developed for side-by-side comparison, and line-drawn images which accompany the original article. In addition to the Brooks and Suchey (1990) method, changes in the morphology of the auricular surface have also been studied. The method was originally developed by Lovejoy et al. (1985) with modification to the method being developed by others (Buckberry & Chamberlain 2002; Igarashi et al. 2005; Mulhern & Jones 2005). Similar to the pubic symphysis, the age assessment on the auricular surface is divided into different 'phases' with each phase given an indication of an age range.

The closure of the sutures of the skull has also been utilized as an age marker. The method requires a visual assessment of the degree of obliteration of the cranial sutures (Key et al. 1994). However, the method has been considered to be unreliable (see Key et al. 1994), but is still yet used amongst some European countries (Márquez-Grant 2015). In addition, tooth formation, eruption and tooth attrition has also been widely used as an age indicator with much of the most important work being developed within the sub discipline of dental anthropology and archaeology (see Hillson 1996; Hillson 2005; Hillson et al. 2006; Hillson 2001). However, dental development and eruption are also dependent on individual variability and are limited in their ability to assess age estimations on sub-adults (Franklin 2010). Additionally, there has been a focus on statistical approaches such as 'Transition Analysis' when estimating age at death. This technique combines morphological data from various skeletal age indicators and adopts a Bayesian approach to calculate a maximum likelihood estimate and confidence interval (Milner & Boldsen 2012).

However, the methods used in age at death estimations have been scrutinised and criticised for being too broad and imprecise, especially when used in forensic investigations (Hoppa & Vaupel 2002). This is due to the fact that very accurate and narrow estimations of an age might be difficult to obtain, but are still expected from legal actors (Márquez-Grant 2015). Studies have also demonstrated that there can be a significant discrepancy between different observers, and issues have also been raised regarding the lack of some appropriate statistical approaches for the estimation of age at death (Lottering et al. 2013; Chen et al. 2011). Concerns have been raised regarding the tendency to overestimate young individuals and underestimate older individuals, when assessing age at death of skeletal remains (Steadman et al. 2006). In addition, the lack of reference samples for people over the age of 60 has created a difficulty in establishing age ranges for older individuals.

In addition to these issues, there remains the fact that various factors (such as nutrition, exercise, diet, socio-economic status, body mass etc.) can influence age markers on the skeleton, which can be difficult to account for when estimating age at death (Merritt 2015; Wescott & Drew 2015). The state of the skeletal preservation could also impact on how many age indicators can be assessed. This is intensified with alterations on the skeleton due to taphonomical changes after death (Márquez-Grant 2015). Additionally, age at death assessments are also highly related to knowing the sex and ancestry, as the methods in age at death estimations will provide different results depending on the sex and ancestry of the skeletal remains. Similarly, methods in age at death assessments are likely to be population based, making some of the methods not appropriate for other populations (Baccino & Schmitt 2006). Additionally, in some cases the police might have an idea of who the remains belongs to, which may result in the forensic anthropologist obtaining that information beforehand (prior to the analysis) and unconsciously being affected by that information resulting in their narrowing down the age estimation to fit the context (Márquez-Grant 2015).

2.3.6. Cognitive Research in Forensic Anthropology

Early research into the possibility of cognitive biases involved in the assessment of skeletal remains was included in a study conducted by Weiss (1972). Here, Weiss compared samples from archaeological skeletal populations and the accuracy of sex estimations on the skeletal sample. Weiss' (1972) results demonstrated a 'male bias' in the assessment of the skeletal remains, with 12% more males than expected when compared to sex ratios in living populations. Weiss argued (after further analysis of the data set) that the flaws in sexing methodologies were more likely to be compounded by bias rather than the population actually containing more males than females. Weiss concluded that this was particularly notable when assessing robust ambiguous skulls as there was a tendency to misidentify ambiguous 'robust' female skulls as males (Weiss 1972). According to Weiss this was due to subtle societal prejudices in the field with regards to robust skeletal skulls 'expected' to be male morphological traits, perhaps resulting in a default male classification.

Equally, Walker (1995) highlighted that there might be a societal prejudice of male and female characteristics (i.e.. females appearing more gracile and males more robust) that could potentially bias the interpretation of archaeological skeletal collections (Walker 1995). For example Walker's (1995) study of the well documented Saint Brides Church skeletal collection in London showed that poorly preserved female pelvises with robust skulls were often misclassified as males. Walker (1995) noticed that the female skulls in the studied population became more robust with age. Walker hypothesised this accounts for the prevalent misidentification of elderly females as males. Furthermore, studies have also identified issues for estimating sex of skeletal remains, where grave goods could potentially cause an expectation bias in sex assessment due to the associated grave artefacts (Effros et al. 2000).

As discussed earlier, there is a rapidly growing response regarding validation studies within the forensic anthropological domain. Acknowledgments of some of the pitfalls and cognitive limitations involved in the traditional techniques applied in the analysis of skeletal remains have been highlighted. Each traditional technique, method, and approach involved in biological profiling has been evaluated in order to establish more standardised methods, especially within visual assessments. While some studies have highlighted parts of cognitive processing that could cause inconsistency and subjective interpretations (such as observer bias, and expectation bias) there is still a lack of empirically based studies that deal with the range of cognitive issues that can occur during forensic anthropological assessments. Studies concerning cognitive issues in judgment and decision-making have gained impetus in the forensic field, and only recently begun to be investigated in the forensic anthropological domain with much of the work being conducted with regards to interpretation issues at the assessment stage of the skeletal remains (Nakhaeizadeh, Hanson, et al. 2014; Nakhaeizadeh, Dror, et al. 2014; Klales & Lesciotto 2016).

2.3.7. Empirical research in Cognitive Bias in Forensic Anthropology

One of the first studies in cognitive biases in forensic anthropology was conducted by Nakhaeizadeh, et al. (2014) concerning the effect of contextual information on the interpretation of skeletal remains. The study involved an experimental design that examined the effect of context on non-metric assessments in sex, ancestry, and age at death estimations. The experiment involved examining the biological profile interpretations of 41 non-novice participants within the field of physical anthropology who assessed the same remains. Each participant was semi-randomly assigned into one of three groups, where two of the groups were given ‘extraneous’ contextual information, before conducting the analysis, with a third group acting as a control with no context provided. The contextual information was provided before establishing a biological profile, and included context that gave indications of sex, origin, and age of the remains (for example given participants result of the DNA analysis). This was audio-recorded and played for each participant prior to the assessment. The study sought to determine if the examiners would be affected by the given context when asked to establish a biological profile. As discussed in previous section 1.2.5, similar studies that examined confirmation bias in other forensic domains have demonstrated that there is a tendency among experts to selectively gather and process information to confirm a hypothesis or preconception by noticing evidence that would validate existing beliefs and expectations. This has been shown to be more powerful and prevalent in ambiguous cases (e.g. Dror, 2011, Found, 2014). The skeletal remains used in this anthropological experiment were of an ambiguous nature. Even though a complete skull and the majority of the postcranial elements were present, the morphological characteristics did not suggest a clear sex or age at the time of death.

The results from the study indicated a statistically significant difference within the assessment of the participants when conducting traditional visual methods on sex, ancestry, and age at death on the skeletal remains. For example, in the assessment of sex, in the group that received contextual information that the remains were female, 100% of the participants concluded the remains to be female. However, in the group that received contextual information that the remains were male, only 14% indicated the remains to be female, 72% indicated the skeletal remains to be from a male, and 14% were undetermined in their conclusion. (see Nakhaeizadeh, Dror et al 2014).

Another study addressing confirmation bias in forensic anthropology was conducted by Klales and Leciotto (2016). Here, the authors explored the idea of confirmation bias and sex estimations of the innominate. The study was conducted on 15 innominates with 7 experienced observers asked to blindly score the three main 3 traits outlined by Phenice (1969). This was done by using a developed 5 scale scoring system, with 1 being female and 5 being male. Each of the 3 traits was scored on a separate day with only the specific trait under examination being visible. Participants were asked after assessing each trait individually to provide an overall impression of the sex as well as scoring each trait again. However, this time all traits were visible and scored simultaneously in combination with examining the whole innominate. The results showed a tendency to change the scaling of single traits on the innominate that have been assessed in isolation, to fit the overall decision reached. The study indicated a confirmation bias where the overall appearance of skeletal elements affected the previous scoring of traits by the participants, conducted in isolation (Klales & Lesciotto 2016).

Furthermore, studies of expectation and motivation bias have also been conducted within forensic anthropology and trauma assessments of skeletal remains. In general terms, trauma interpretations involve descriptive analyses of fracture morphology and modifications on bones (Symes et al. 2012; Passalacqua & Fenton 2012; Blau 2016). It has been recognised that accuracy in trauma assessments within forensic anthropology are highly related to training and experience of the expert within the domain (Pinheiro et al. 2015; Blau 2016). This is due to the complexity of the trauma that could be involved in the interpretation process (Symes et al. 2012). An preliminary study was undertaken to assess the degree to which the expectations and interpretations of trauma would vary depending on the context provided (Nakhaeizadeh, Hanson, et al. 2014). The study demonstrates that cognitive interpretation issues are also apparent in the visual assessment of trauma analyses. Three different websites were created, each with fourteen identical images with different levels of trauma traits on skeletal remains. Ninety-nine participants with a physical anthropological background were then asked to assess each image. Each website was associated with different contextual information, with one context indicating a high trauma expectation (mass grave setting), and the other context indicating low trauma expectation (archaeological setting). The results of the study indicated a higher scoring of trauma identification responses among participants assessing images in a high trauma context setting, compared to participants evaluating the same pictures in a setting with low trauma context.

For example, image number six on the websites illustrated a foramen (an opening/hole) manifested in the distal end of the sternum (sternal aperture). This foramen is well known among anthropologists to manifest itself on the sternum as a biological variation. Participants in the archaeological website setting distinguished the trait not to be associated with trauma with 94% of the participants selecting the “no trauma” option. This was also distinguished in the control group website, where

81% of the participants selected the "no trauma" option. However, there was a distinct difference in the interpretation of the trait in the mass grave website. Only 33% of the participants from the mass grave website setting selected the feature to be "no trauma," with the majority of the participants selected the feature to be of "possible trauma." Similar results were found in four other comparable images (See Nakhaeizadeh, Hanson, et al. 2014).

2.3.8. Future directions

The issue of cognitive bias and its effects in disciplines undertaking forensic investigations has been increasingly discussed and described in the published literature. It is clear that cognitive biases are not limited to a specific 'domain' but can be manifested within any discipline involving human decision-making and interpretations (e.g. Dror and Fraser-Mackenzie, 2008). In the forensic domain it has been empirically established that cognitive biases may influence data collection, analysis, interpretation and review of conclusions, specifically in the use of subjective methods (e.g. Kassin, *et al.,* 2013) . In forensic anthropology, the majority of methods used are subjective by their nature, since they are heavily based on human judgment by visual observations. Although a good number of studies have been conducted upon methodological 'error' rates and inter-intra observer variability (as previously discussed in this section of this chapter), the cognitive impacts at work during the assessment of human remains has only recently begun to be assessed in the published literature. Despite the fact that contextual and environmental effects have been shown to be powerful influences on how people construct and interpret information, there is still much work needed within the forensic anthropological domain in order to fully understand cognitive and contextual biases in the positive identification of skeletal remains, from crime scene to court. This is ever more important at this stage due to the fact that the issues regarding the admissibility of evidence and expert witness testimonies have been consistently raised in the validation of methods used by forensic scientists. This also includes acknowledging and accounting for cognitive factors that are intrinsically a part of forensic reconstruction approaches.

The body of knowledge concerning the application of decision theory to forensic anthropology could therefore, be of considerable benefit to the forensic community.

Conducting further research into cognitive biases and forensic anthropology, will produce data that will test empirically when such factors are more likely to occur and equally when they do not. Empirical research will aid in the development of an evidence base that addresses the cognitive constraints on non-metric assessments, (which are commonly used in the teaching and practice of forensic anthropology) and identifying the settings when cognitive issues are more or less likely to impact an identification process. This will not only strengthen the discipline itself but will also empower the discipline to lead the way in the development of the broad range of forensic sciences. Therefore, to tackle the challenge of combining and interpreting different sources of information, and achieving transparency in decision-making and evidence-based conclusions, a better understanding is required within the field of forensic anthropology of the underlying process of decisions being made and potential cognitive influences in the interpretation of skeletal remains throughout the forensic science framework.

2.4. Aims and Objectives

The aim of this research is to further examine the extent to which cognitive biases are present within forensic anthropological methods. This research will specifically seek to understand the degree of contextual effects in forensic anthropological assessments and present ways that can mitigate the impacts in biological profiling. The research will focus upon what methodological procedures and phases within forensic anthropology and the establishment of a biological profile are more or less vulnerable to cognitive interpretation issues and under what circumstances. This will be achieved by undertaking experiments to test for cognitive and contextual effects empirically within forensic anthropological procedures and methods. More specifically the objectives are:

1. To examine the effect of contextual information on judgment and decision-making in forensic anthropological visual methods used in the establishment of a biological profile. This will specifically address whether contextual information can affect previous judgments when assessing skeletal remains of an ambiguous nature (Chapter 3 Experiment 1 and 2)

2. To investigate the potential effects of initial exposure to context at a crime scene upon judgment and decision-making in the subsequent assessment of skeletal remains. This will specifically examine whether early exposure to 'extraneous' contexts in the excavation of skeletal remains cascade and thereby affect the subsequent assessment of the skeletal analysis (Chapter 4 Experiment 3)

3. To address whether the order of examination of skeletal remains influences, a) the interpretation of the subsequent skeletal element, and b) act as an influence and determine the final conclusion of the assessment (Chapter 5 Experiment 4).

The research undertaken allowed for a holistic examination of the stages and methodological procedures when and to what extent cognitive factors may affect performances and render the judgments of forensic anthropologists to be compromised and equally when they do not.

Chapter 3. Cognitive and Contextual Influences in Forensic Anthropology: the implication of observer effect in biological profiling (Experiment 1 and 2)

3.1. Introduction

As outlined in chapter 2 section 2.1 a substantial body of research in cognitive psychology and decision-making has identified that assumptions, concepts, beliefs, and information retrieved from memory, form a mind-set that guides perception and processing of new information (Kassin et al. 2013). It has been recognised that human cognition employs simplified information processing strategies to ease the load of mentally processing information when making judgments and decisions (Anderson & Kellam 1992). Over the years different types and precursors of cognitive biases have been recognised, (see Chapter 2.1 section 2.1.5 for further details) such as observer effect (Cooley & Turvey 2011) and contextual influences (Fraser-Mackenzie et al. 2013). For example, context effect may occur when forensic examiners are affected by the contextual information available to the analyst prior to their assessment, or by the context of the crime (Saks et al. 2003). Context effect is highly related to observer effect (when the result of an observation in a specific circumstance is affected by the observer), such that the preconceptions of an observer and sometimes their motives, are thought to influence the interpretation of evidence (Krane et al. 2008).

Both observer effect and contextual biases have been established empirically to affect the decision-making of forensic scientists. As outlined in Chapter 2.2 section 2.2.7 studies within the fingerprint domain and DNA demonstrated an inconsistency in analysis, as well as biased interpretations, when experts were subjected to different types of contextual information (Dror & Hampikian 2011; Dror et al. 2006; Dror et al. 2005; Dror et al. 2011). It is clear that in a similar manner, both forensic anthropologists and physical anthropologists/archaeologists could potentially be exposed to contextual information in terms of case files, gravesite descriptions as well as previous osteological reports. Therefore, understanding how contextual influences could potentially affect the interpretation of skeletal remains needs to be further established. In addition, in a manner akin to other forensic disciplines, some of the methods used in forensic anthropology have been argued to be limited because of their subjective nature (Walrath et al. 2004; Márquez-Grant 2015). As highlighted in Chapter 2 section 2.3 this has specifically been highlighted within non-metric assessments where it has been discussed that the techniques used are generally reliant upon observation and the specialised experience of the observer (Hefner et al. 2007) possibly making them vulnerable to subjective interpretations and contextual influences. For example, as outlined in Chapter 2 section 2.3.6, early study by Weiss (1972) found that there is a tendency to misidentify ambiguous robust male skulls as females. Equally, Walker (1995) highlighted that there might be a societal prejudice of male and female characteristics that could potentially bias the interpretation of archaeological skeletal remains. Furthermore, studies have also identified issues for estimating sex of skeletal remains, where grave goods could potentially cause an expectation bias in sex assessment due to the associated grave artifacts (Effros et al. 2000).

This chapter will therefore add to the forensic anthropological literature by examining the effect of contextual information on judgment and decision-making in forensic anthropological methods. The experiments presented here specifically address whether contextual information can affect previous judgments when assessing skeletal remains of an ambiguous nature. A series of experiments was designed to assess the consistency in the interpretations of skeletal remains made at different times under different contexts. The experiments were carried out in order to gain insights into whether the decisions of participants were independent and thus consistent regardless of extraneous influences, or alternatively whether the participants changed their previous decisions as a result of being given context.

3.2. Methodology

The experimental study was designed to investigate the effect of contextual information in biological profiling, when decisions are made on ambiguous skeletal remains. This was achieved through two experimental designs with a pilot study (Experiment 1) and a subsequent study that built on the findings of the pilot study (Experiment 2).

3.2.1. Experimental Design

3.2.1.1. Experiment 1 (Pilot)

Participants were asked to establish sex, ancestry, and age at death estimations (the most common attributes to provide a biological profile) on two different occasions on a set of skulls and os coxa, based on visual assessments. This approach was undertaken to firstly establish a baseline control. This was done in order to assess what decisions would be made without any contextual information. Secondly to provide contextual information (prior to participants assessments) to assess whether the interpretation of the skeletal remains made by the participants would be differ and change when given 'contradicting' context. The experiment took place over a two-month period with a 3-week interval in-between each phase of the study (see Table 3.1). The timeframe of the experiment was based on the availability of the participants as well as time constrains.

3.2.1.2. Experiment 2 (Follow up study)

Based on the limitations of the pilot study (Experiment 1), Experiment 2 was designed as a follow up and control study. This was done in order to a) address the constraints observed in Experiment 1 (such as the lack of a second control-reliability round) in order to cap any variables that could have affected the results, and b) to replicate the study in order to increase the number of participants. Participants in Experiment 2 were also asked to establish sex, ancestry and age at death estimations on three different occasions on the same a set of skulls and os coxa used in Experimental 1, based on visual assessments. Similar to Experimental 1, this approach sought to firstly set a baseline control (in order to assess with a control condition what decisions would be made without any contextual information,) secondly, to provide contextual information (to assess whether the decisions of participants would be different and change when given contradicting context), and thirdly, to again provide participants with no contextual information (for a second time) as a further control to test for reliability and exclude any 'noise ' that might have affected the results from Experiment 1. In addition, participants in Experiment 2 were also asked to assess their confidence level for each assessment on a numerical scale from 1-100. This was added in the experimental procedure in order to understand the decision-making process and the confidence in the decision-making for the participants. The experiment took place over a 3-month period with a 3 -week interval in-between each phase of the study (see Table 3.1).

Table 3.1 Time table of the research design

	Week	Circumstances of the analysis	Number of participants
Experiment phase (Experiment 1)			
Analysis 1 (baseline – control)	1	No contextual information	18
Analysis 2 (context)	4	Contradictory contextual information	18
Experiment phase (Experiment 2)			
Analysis 1 (baseline – control)	1	No contextual information	22
Analysis 2 (context)	4	Contradictory contextual information	22
Analysis 3 (reliability-control	7	No contextual information	22

3.2.2. Materials

Three skulls and two os coxa were selected from the skeletal collection curated by the UCL Institute of Archaeology, with the material originating from archaeological excavations. The selection of the skeletal remains was identical for both Experiment 1 and 2. The selection was made by undertaking a pilot test run to assess the characteristics of each set of remains to determine the degree of ambiguity, thereby identifying remains where the morphological features present on the skulls and the os coxa were of a complex nature, showing mixed traits of female, or male characteristics. Furthermore, no extreme trait distinction regarding ancestry was included as well as ambiguity in age at death, due to the wide age ranges generally assessed in forensic anthropology. This was of specific importance in this study because cognitive influences have been shown to be more predominant in ambiguous and complex cases (Kassin et al. 2013).

The skulls were intact and in a good condition, making it possible to conduct a visual assessment on sex and ancestry. Equally, the two os coxa were both partly intact with the *Ilium*, *Pubic* and *Ischium* bones presented for visual analysis on sex and age at death estimations (see Fig. 3.1).

Figure 3.1 Showing the three skulls and two os coxa skeletal used for this study

3.2.3. The contextual information

The contextual information was provided separately for each component of the skeletal remains, and was presented next to each skeletal element in the form of a short report. The information included elements of grave context descriptions as well as information taken from the osteological report of the site, as exemplified in the extract below.

> "This individual was excavated from St Brides Crypt, known to represent affluent individuals, with the majority of the burials being of Caucasian descent. According to grave context descriptions, a whalebone (commonly used for corset stays) was found in addition to the skeletal remains"

The nature of the contextual information used in this study was based upon archaeological and osteological reports and archives, grave artefacts, and site descriptions. This was done to provide the participants with realistic and credible contexts for each skeletal element. Consequently, information regarding a forensic setting was not included in this study and equally, each context for each skeletal element differed but all indicative of a certain sex, ancestry or age at death (see Appendix B for full list over the contextual information for each skeletal element). The contextual information given to the participants contradicted the majority of the previous estimations of the participants. For example, if a participant estimated the skeletal remains to be male in previous decisions, context was given to indicate a female and vice versa. A similar approach was taken for ancestry and age at death estimations.

3.2.4. Participants

3.2.4.1. Experiment 1

A total number of 18 forensic anthropology Masters students completed the study. The participants came from various backgrounds (most having undertaken a undergraduate qualification in physical anthropology), but all had the appropriate training in biological profiling in order to complete the exercise. The students were not informed of the true nature of the study as that would have affected their decision making process. Instead the experimental design was hidden in a series of practical workshop assignments, organised for the students to practice their knowledge and skills in forensic anthropology. The workshop included other non-ambiguous skeletal elements in which the students were asked to assess. The five ambiguous skeletal elements pertinent to this study were all placed together with the non- ambiguous elements for the participants to evaluate. All participants were informed of the true nature of the study after the experiment was completed following standard ethical code of practice (see Appendix B for Consent form).

3.2.4.2. Experiment 2

Experiment 2 was conducted in identical fashion to Experiment 1 with the exception of the number of participants, which increased from 18 to 22.

3.2.5. Procedure

Experimental 1 and 2 both followed the same standard and procedure. Each skeletal element was laid out separately, creating different sections and stations for participants to conduct their assessment. Participants were divided into two groups, mainly because of space, and the design of the workshops. This limited the ability to customise the contextual information to fit each participant specifically. Therefore, the contextual information given to each skeletal element was created to contradict the answer given by the majority of students in both groups based upon the baseline results. Consequently, this resulted in some of the students receiving contextual information that supported their initial judgments. Each participant was directed to an individual station where instructions were given. Participants had five minutes at each station before asked to move to the following one. The time limit was set based upon the pilot test run conducted during the experimental design phase, as well as the time limit provided for participants for when doing forensic anthropological pop-up exams. The order of appearance of the ambiguous skeletons was changed for each of the rounds during the 3-month period. In addition, participants were not aware that they were assessing the same skeletal remains during the different occasions.

Participants were given access to reference materials for sex, ancestry and age at death estimations. This list was compiled from the most common historically traditional non-metric methods used and studied within the field of physical anthropology and osteological techniques, focused on methods from Phenice (1969), Bukistra and Uberlaker (1994), Rhine (1990), and Suchey-Brooks (Brooks & Suchey 1990). Due to the experiment being hidden in a practical workshop for the students, participants were informed that they were going to conduct a biological profile exercise in from of pop up test as part of the workshop. All instructions were given orally in which students were asked to follow the biological profile form/answering sheet given to them at the start of the workshop. The form was previously developed for the students to use in class an in the forensic anthropological workshop, and followed a similar order and point scoring to the practical mini test. By using the form and answering sheet from the practical workshop ensured that the students were conducting assessments and following procedures that they were already familiarised with. The form included assessing sex, ancestry, and age at death, using visual methods as well as assessing trauma, taphonomy and pathology for some skeletal elements. Each section was sub-divided with all morphological traits outlined for each assessment and methods, in order to make it easier for the participants to assess the skeletons based on the time limit provided. However, only the answers for the five skeletal elements used for this study (which was hidden in the workshop exercise) was analysed from the form (see Appendix B for answering sheet for the five skeletal elements used). Furthermore, participants were asked to fill in their final answer based on the categorical options given for each assessment. For example, in sex estimation the final categorical variables given were 'Female', 'Male', or 'Undetermined' (similar to categories used within traditional methods). Participants were allowed at all times to conclude 'undetermined' as an option at every stage of the analysis. Participants were also given extra time (at the end of the session) and line space to describe and justify their answer (if they wanted too) in order to better understand the decision-making process.

Furthermore, in order to avoid any biases entering the process, the author was not present during the experimental procedures. As mentioned earlier, the experiment design was hidden in a series of practical workshops, part of the forensic anthropological course taken by the participants. Therefore, the principle lecturer in forensic anthropology was present during the workshops and the experiments in order to avoid any suspicion that may have arisen from deviations from the weekly routines that participants were familiar with. This also minimized any personal expectations and contextual biases that could have been introduced by the author during the procedure due to the authors' prior knowledge of the true nature of the study.

3.2.6. Analysis

3.2.6.1. Experiment 1

Descriptive statistics and percentage analysis was used to analyse and present the data for Experiment 1

3.2.6.2. Experiment 2

Descriptive statistics and percentage analysis was also used to present the data for Experiment 2. In addition, a series of Chi-square was conducted to determine statistically whether the distribution of categorical variables between each group (Baseline, Context, Control) for each skeletal element (skull 1, skull 2, skull 3, os coxa 1, os coxa 2) and the assessments (sex, ancestry and age at death) differed significantly from one and other as a function of the contextual information. Chi-square was first used to establish if there was a significant different between Baseline vs. Reliability-Control, and then furthermore applied to distinguish if there was a significant difference between the consistent answers from participants in Baseline/Control vs. Context. To compare the confidence level of participants for all three test-runs, a one way repeated measure ANOVA was used as well as series of post hoc tests in form of paired t-tests.

3.3. Results

3.3.1. Experiment 1

A total number of 18 participants took part in the study, completing the two phases (Baseline Control, Context), resulting in a total amount of 360 decisions, across all skeletal elements with 180 decisions made in each phase (see Table 3.2 for full detail). Out of the 180-paired decisions made (comparing Baseline Control with Context,) 46 were not testable due to incomplete answers or the context being the same as their initial answers, resulting in a comparison of 134 decisions between Baseline and Context. Out of the 134 decisions made by participants, 103 (76.8%) were affected in their decision-making when given context, with only 31 decisions (23%) not being affected by the context provided in the second phase of the study.

Table 3.2 showing the distribution of all decisions being made for each skeletal element

Decision	Skull 1	Skull 2	Skull 3	os coxa 1	os coxa 2	Total
Baseline control						180
Sex	18	18	18	18	18	
Ancestry	18	18	18	**x**	**x**	
Age at Death	**x**	**x**	**x**	18	18	
Context						180
Sex	18	18	18	18	18	
Ancestry	18	18	18	**x**	**x**	
Age at Death	**x**	**x**	**x**	18	18	
Total						**360**

3.3.1.1. Overall Distribution

Figure 3.2 demonstrates the overall distribution of the effect of context on the participants for each skeletal component, with Table 3.3 giving descriptive information on each category.

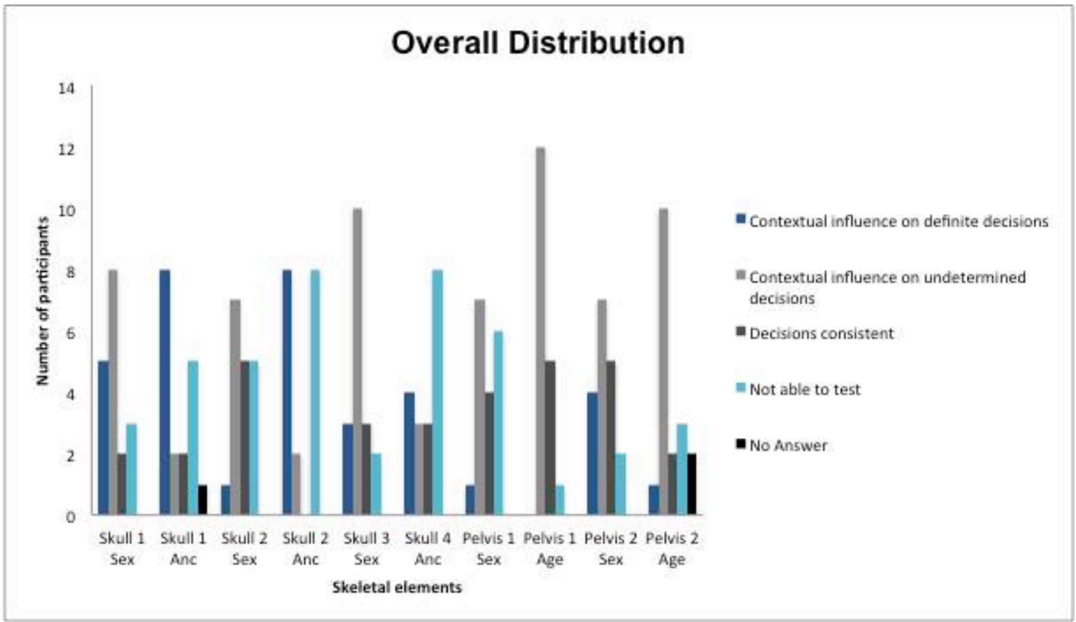

Figure 3.2 Demonstrating the overall distribution of the effect of context on sex, ancestry (Anc) and Age at death (Age) assessment of the skeletal elements.

Table 3.3 showing descriptive information for each decision category

Contextual influence on definite decisions	Contextual influence on undetermined decisions	Decisions consistent	Not able to test	No Answer
Participants that on previous decisions did not estimate the skeletal remains to be 'undetermined' but made a an definite/categorical estimation on sex, ancestry or age at death on previous decisions of the skeletal remains, but did however change their decision-making when context was given, resulting in agreeing with the context	Participants that were undetermined in their decision-making process on previous decisions but however, confirmed with the contextual information given, resulting in a change on previous interpretations of the skeletal remains.	Participants that were not affected by the context regardless of previous decision-making.	Participants within the groups that were given the same context as previous decisions	Participants that did not complete their final answer

3.3.1.2. *Skull 1 Sex Assessment*

The total amount of participants completing a sex assessment on previous decisions on skull 1 was fifteen. The results indicate two participants (13.3%) were not affected by the contextual information on previous decisions and thirteen participants (86.7%) confirming with the context, resulting in a change of previous decisions (see Figure 3.2 (Skull 1 sex) for further breakdown of the effect of context on participants interpretation of the skeletal element).

3.3.1.3. *Skull 1 Ancestry Assessment*

The total amount of participants completing an ancestry assessment on previous decisions on skull 1 was twelve. The results indicate two participants (16.7%) were not affected by the contextual information on previous decisions and ten participants (83.3%) confirmed the context, resulting in a change of previous decisions (see Figure 3.2 (Skull 1 ancestry) for further breakdown of the effect of context on participants interpretation of the skeletal element).

3.3.1.4. *Skull 2 Sex Assessment*

The total amount of participants conducting a sex assessment on previous decisions on skull 2 was thirteen. The results indicates, five participants (38.5%) not being affected by the contextual information on previous decisions and eight participants (61.5%) confirming with the context, resulting in a change of previous decisions (see Figure 3.2 (Skull 2 sex) for further breakdown of the effect of context on participants interpretation of the skeletal element).

3.3.1.5. *Skull 2 Ancestry Assessment*

The total amount of participants conducting an ancestry assessment on previous decisions on skull 2 was ten. The results indicates, all ten participants (100%) to be affected by the contextual information confirming with the context, resulting in a change of previous decisions (see Figure 3.2 (Skull 2 ancestry) for further breakdown of the effect of context on participants interpretation of the skeletal element).

3.3.1.6. Skull 3 Sex Assessment

The total amount of participants conducting a sex assessment on previous decisions on skull 3 was sixteen. The results indicates, three participants (18.8%) not being affected by the contextual information on previous decisions and thirteen participants (81.2%) confirming with the context, resulting in a change of previous decisions(see Figure 3.2 (Skull 3 sex) for further breakdown of the effect of context on participants interpretation of the skeletal element).

3.3.1.7. Skull 3 Ancestry Assessment

The total amount of participants conducting an ancestry assessment on previous decisions on skull 3 was ten. The results indicates, three participants (30%) not being affected by the contextual information on previous decisions and seven participants (70.0%) confirming with the context, resulting in a change of previous decisions (see Figure 3.2 (Skull 3 ancestry) for further breakdown of the effect of context on participants interpretation of the skeletal element).

3.3.1.8. Os coxa 1 Sex Assessment

The total amount of participants conducting a sex assessment on previous decisions on os coxa 1 was twelve. The results indicates, four participants (33.3%) not being affected by the contextual information on previous decisions and eight participants (66.7%) confirming with the context, resulting in a change of previous decisions (see Figure 3.2 (os coxa 1 sex) for further breakdown of the effect of context on participants interpretation of the skeletal element).

3.3.1.9. Os coxa 1 Age at Death Assessment

The total amount of participants conducting an age at death assessment on previous decisions on os coxa 1 was seventeen. The results indicates, five participants (29.4%) not being affected by the contextual information on previous decisions and twelve participants (70.6%) confirming with the context, resulting in a change of previous decisions (see Figure 3.2 (os coxa 1 age) for further breakdown of the effect of context on participants interpretation of the skeletal element).

3.3.1.10. Os coxa 2 Sex Assessment

The total amount of participants conducting a sex assessment on previous decisions on os coxa 2 was sixteen. The results indicates, five participants (31.3%) not being affected by the contextual information on previous decisions and eleven participants (68.8%) confirming with the context, resulting in a change of previous decisions (see Figure 3.2 (os coxa 2 sex) for further breakdown of the effect of context on participants interpretation of the skeletal element).

3.3.1.11. Os coxa 2 Age at Death Assessment

The total amount of participants conducting an age at death assessment on previous decisions on os coxa 2 was thirteen. The results indicates, two participants (15.4%) not being affected by the contextual information on previous decisions and eleven participants (84.6%) confirming with the context, resulting in a change of previous decisions (see Figure 3.2 (os coxa 2 age) for further breakdown of the effect of context on participants interpretation of the skeletal element).

3.3.2. Experiment 2

3.3.2.1. Overall analysis of the decision making of participants

A total number of 22 participants took part in the study, completing all 3 phases (Baseline Control, Context and Reliability Control), resulting in a total amount of 660 decisions, with 220 decisions made in each phase (see Table 3.4).

Examining reliability, out of the 220 decisions made by participants in the Baseline control phase, 170 decisions (77%) were consistent with the decision made when compared to the Reliability Control phase, with only 50 decisions (23%) not being consistent with previous decisions. Such reliability in expert performance addresses an aspect in expert decision making not related to ‘biasability’ (see Level 5 in the Hierarchy of Expert Performance (HEP), in contrast to the ‘biasability’ aspect, Level 7 –Dror, 2016).

Table 3.4 showing the distribution of all decisions being made for each skeletal element

Decision	Skull 1	Skull 2	Skull 3	os coxa 1	os coxa 2	Total
Baseline control						220
Sex	22	22	22	22	22	
Ancestry	22	22	22	x	x	
Age at Death	x	x	x	22	22	
Context						220
Sex	22	22	22	22	22	
Ancestry	22	22	22	x	x	
Age at Death	x	x	x	22	22	
Reliability Control						220
Sex	22	22	22	22	22	
Ancestry	22	22	22	x	x	
Age at Death	x	x	x	22	22	
Total						**660**

3.3.2.2. Descriptive statistics of the overall effect of context on the consistent decisions

Out of the 170 consistent decisions made by participants across all skeletal elements, 107 (63%) were changed to align with the context given. Further, only 44 decisions (26%) were not affected by the context provided in the second phase of the study. In addition a total of 19 of the decisions (11%) made by participants were not used, due to the context given being the same as participants answer from both control phases.

3.3.2.3. Descriptive statistics and Chi-Square of Skull 1

Sex assessment

A total number of 16 participants were consistent in their decisions when comparing Baseline Control vs. Reliability Control. Out of the 16 consistent participants, 13 participants (81%) were affected by the contextual information, changing previous decisions with only 3 participants (19%) not being affected by the contextual information (see Figure 3.3 and 3.4 for further details on the effect of female and male context). A Chi-square test comparing the 16 consistent participant decision-making from Baseline/Control vs. Context, revealed a significant difference for participants receiving female context with a p-value <0.05. However, no significant difference was identified for participants receiving male context showing a p-value >0.05 (see Table 3.5).

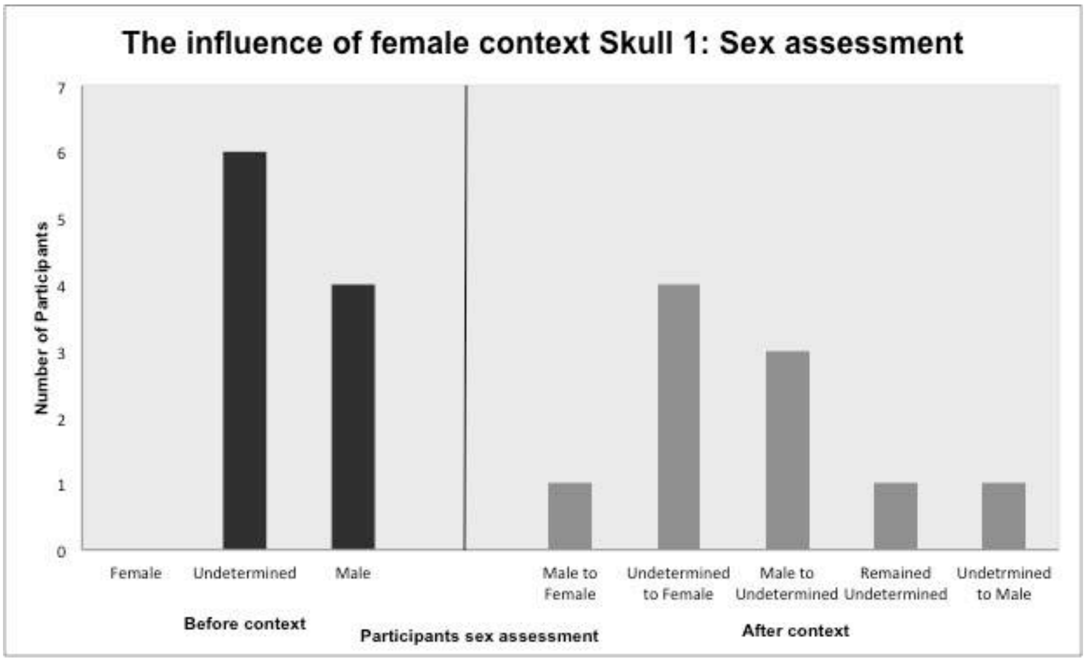

Figure 3.3 showing the interpretations of the skeletal remains before and after receiving female context, demonstrating a change in the interpretations (to align with the context) for eight participants.

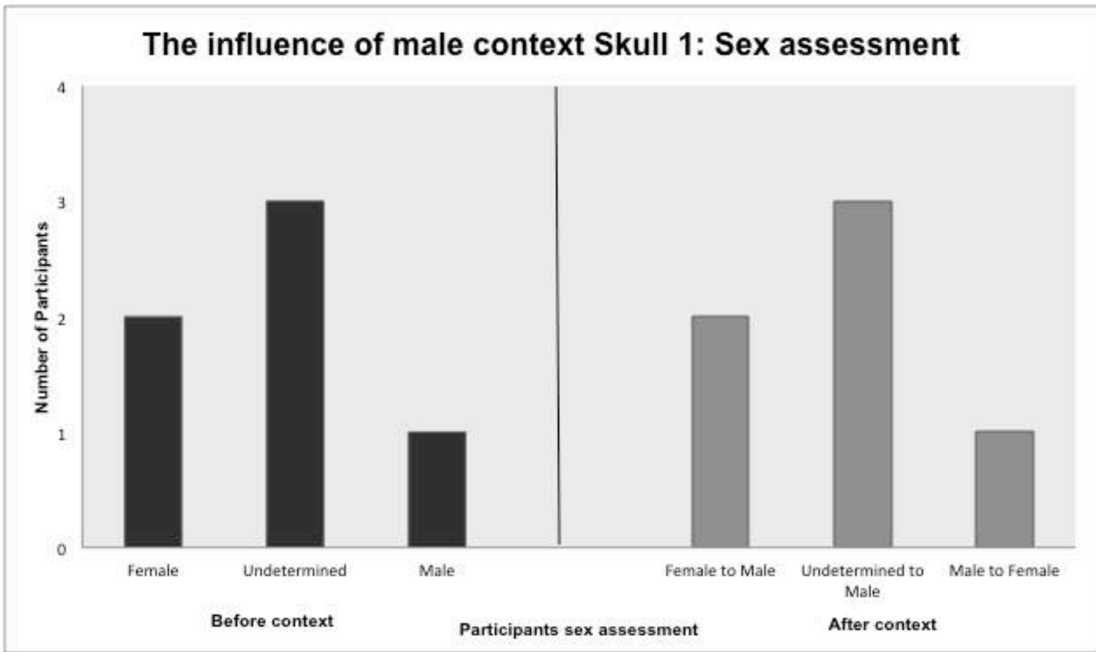

Figure 3.4 showing the interpretations of the skeletal remains before and after receiving male context, demonstrating a change in the interpretations (to align with the context) for five participants.

Ancestry assessment

A total number of 16 participants were consistent with their previous decisions when comparing Baseline Control vs. Reliability Control. Of the 16 consistent participants, 9 participants (56%) were affected by the contextual information, changing previous decisions, with only 4 participants (25%) not being affected by the context. A total number of 3 participants (19%) were untestable, as the contextual information given to those participants reinforced their answers given in the no context conditions, and therefore not comparable (see Fig. 3.5 and 3.6 for further details on the effect of White and Asian context) The 3 untestable participants were taken out of the data set for the Chi-square test when comparing Baseline/Reliability Control vs. Context. The results from the Chi-square test revealed a significant difference for participants given Asian context with a p value of <0.05. No significant difference was found for participants given White contextual information showing a p-value >0.05 (see Table 3.5).

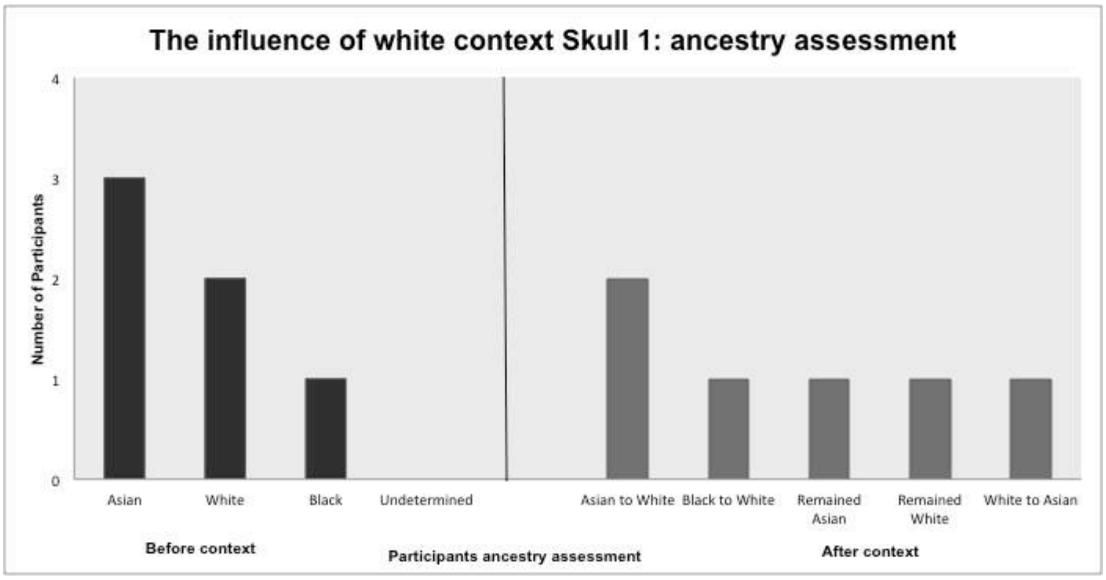

Figure 3.5 showing the interpretations of the skeletal remains before and after receiving white context, demonstrating a change in the interpretations (to align with the context) for three participants.

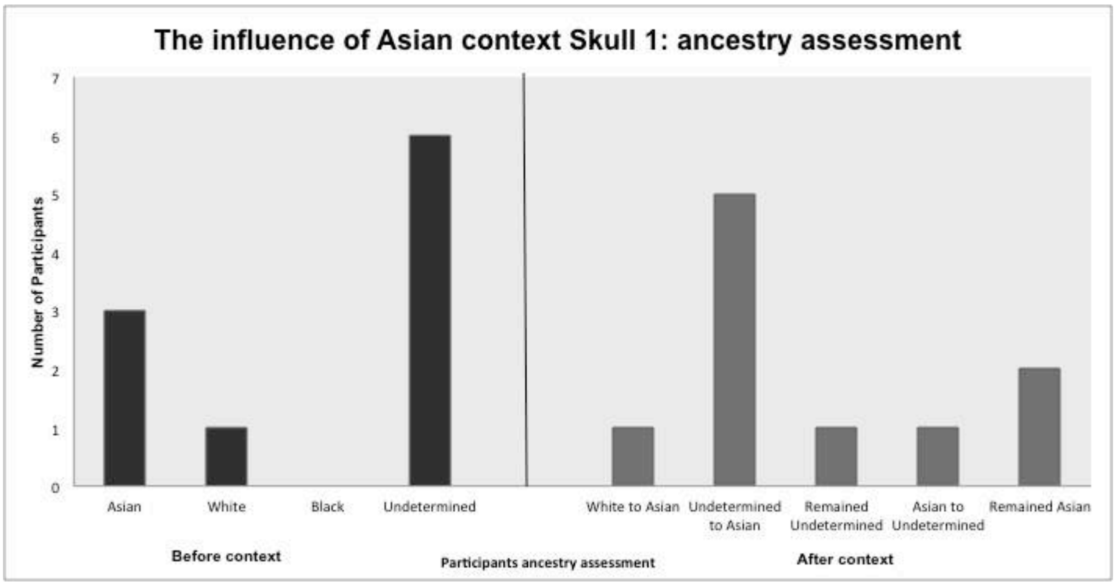

Figure 3.6 showing the interpretations of the skeletal remains before and after receiving Asian context, demonstrating a change in the interpretations (to align with the context) for six participants

Sex assessment

A total number of 16 participants were consistent with their previous decisions when comparing Baseline Control vs. Reliability Control. Of the 16 consistent participants, 10 participants (62%) were affected by the contextual information, changing previous decisions, with 3 participants not being affected by the context (19%) and 3 participants being untestable (19%) (see Fig. 3.7 and 3.8 for further details on the effect of female and male context). The 3 untestable participants were removed from the data set to conduct a chi-square test to compare Baseline/Reliability Control vs. Context. The results from the Chi-square test revealed a significant difference for participants given male context with a p-value < 0.01 and no significant difference for participants receiving female context with a p-value >0.05 (Table 3.5).

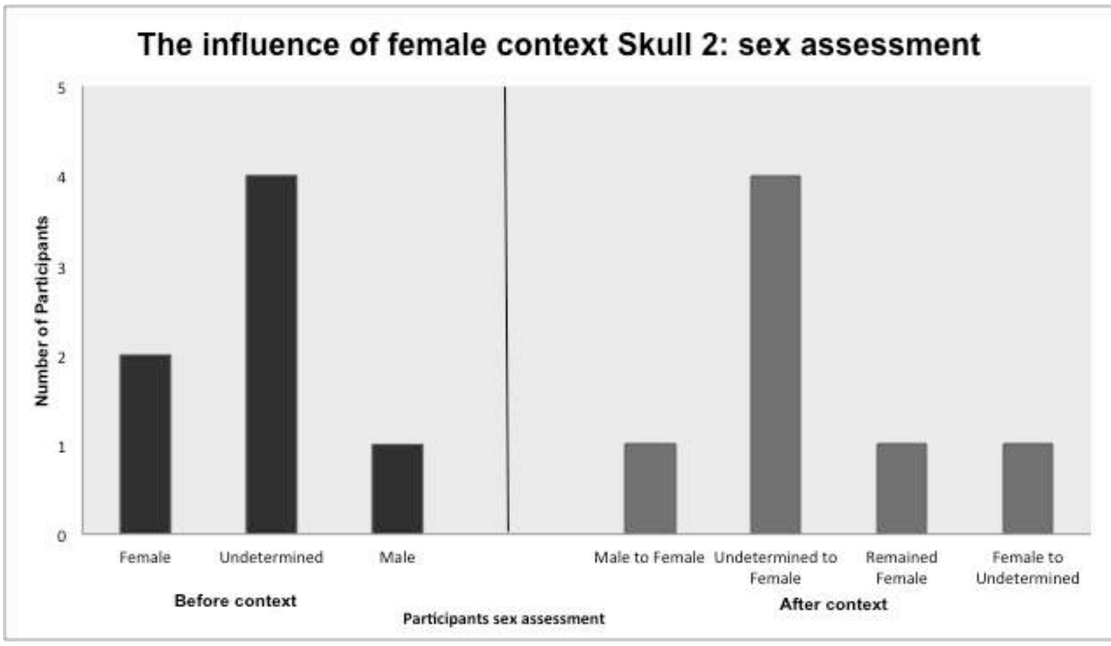

Figure 3.7 showing the interpretations of the skeletal remains before and after receiving female context, demonstrating a change in the interpretations (to align with the context) for five of the participants.

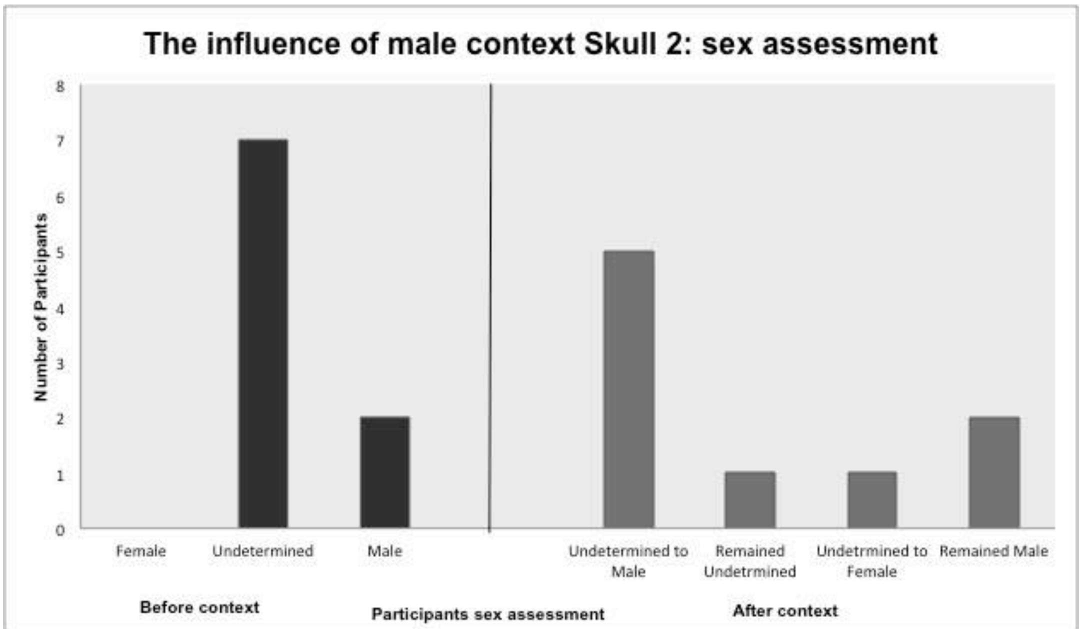

Figure 3.8 showing the interpretations of the skeletal remains before and after receiving male context, demonstrating a change in the interpretations (to align with the context) for five participants.

Ancestry assessment

A total number of 19 participants were consistent with their previous decisions when comparing Baseline Control vs. Reliability Control. Out of the 19 consistent participants, 12 participants (63%) were affected by the contextual information, changing previous decisions, with 2 participants not being affected by the contextual information (11%). A total number of 5 participants were untestable (26%) (see Figure 3.9 for further details). The 5 untestable participants were taken out for the Chi-square test when comparing Baseline/Reliability Control vs. Context The results from the Chi-square test revealed a significant difference between the decisions made concerning ancestry with and without context with a p-value <0.001. See Table 3.5 for further details.

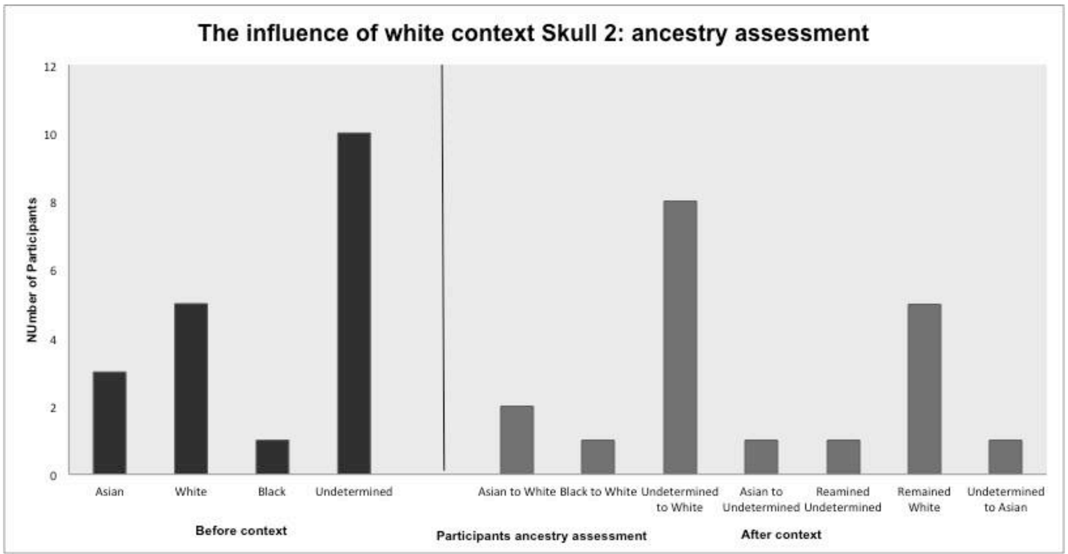

Figure 3.9 showing interpretations of the skeletal remains before and after receiving White context, demonstrating a change in the interpretations (to align with the context) for twelve participants.

3.3.2.4. Descriptive statistics and Chi-Square of Skull 3

Sex assessment

A total number of 18 participants were consistent with their previous decisions when comparing Baseline Control vs. Reliability Control. Out of the 18 consistent participants, 11 participants (62%) were affected by the contextual information, changing previews decisions, with 5 participants not being affected by the context (29%), and 2 participants being untestable (11%) (see Figure 3.10 and 3.11 for further details on the effect of female and male context). The 2 untestable participants were taken out for the Chi-square test when comparing Baseline/ Reliability Control vs. Context. The results from the Chi-square test revealed a significant difference for participants given female context with a p-value <0.01. No significant difference was found for participants given male context showing a p-value >0.05. See Table 3.5 for further details.

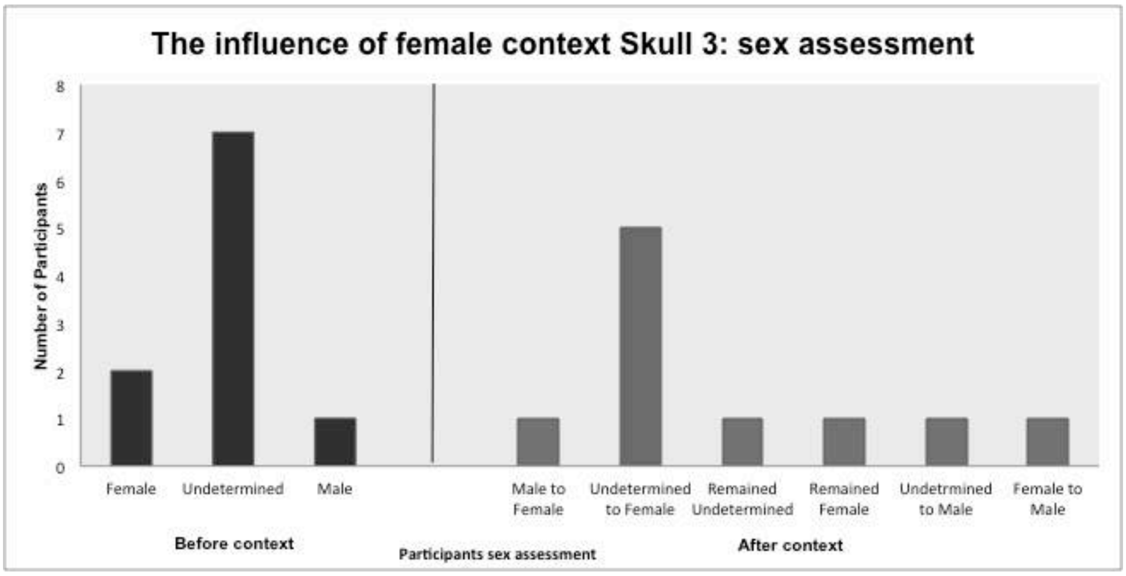

Figure 3.10 showing the interpretations of the skeletal remains before and after receiving female context, demonstrating a change in the interpretations (to align with the context) for six participants.

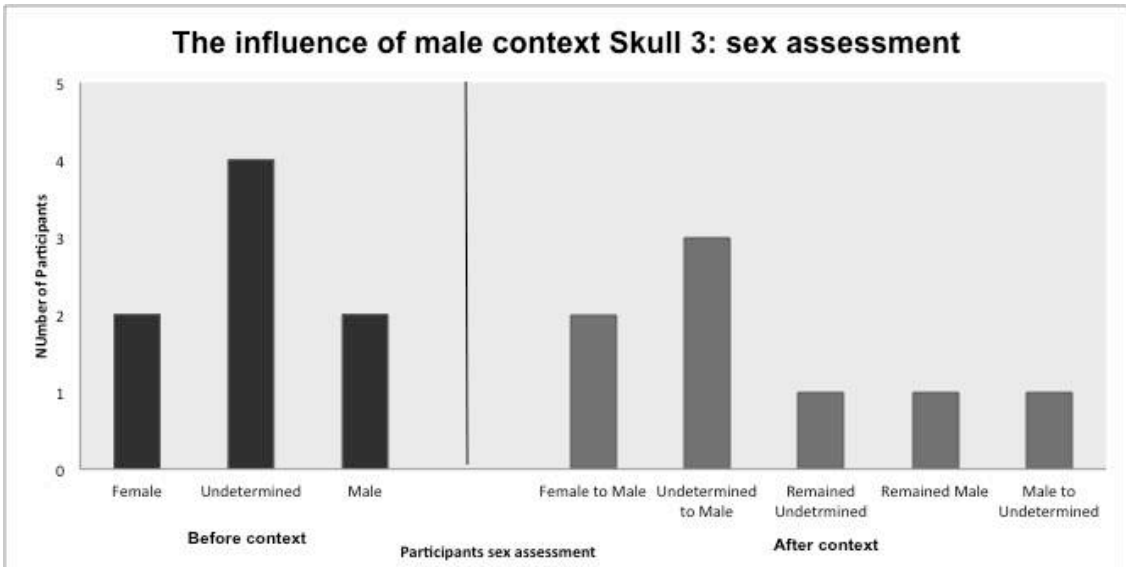

Figure 3.11 showing the interpretations of the skeletal remains before and after receiving male context, demonstrating a change in the interpretations (to align with the context) for five participants.

Ancestry assessment

A total amount of 17 participants were consistent with their previous decisions when comparing Baseline Control vs. Reliability Control. Out of the 17 consistent participants, 13 participants (76%) were affected by the contextual information, changing previous decisions, with 2 participants not being affected by the contextual information (11%), and 2 participants being untestable (11%) (see Figure 3.12 and 3.13 for further details on the effect of White and Asian context). The 2 untestable participant was taken out for the Chi-square test when comparing Baseline/Control vs. Context. The results from the Chi square test revealed a significant difference for participants given White context with a p-value <0.05. A significant difference was also found for participants given Asian context showing a p-value <0.01. See Table 3.5 for further details

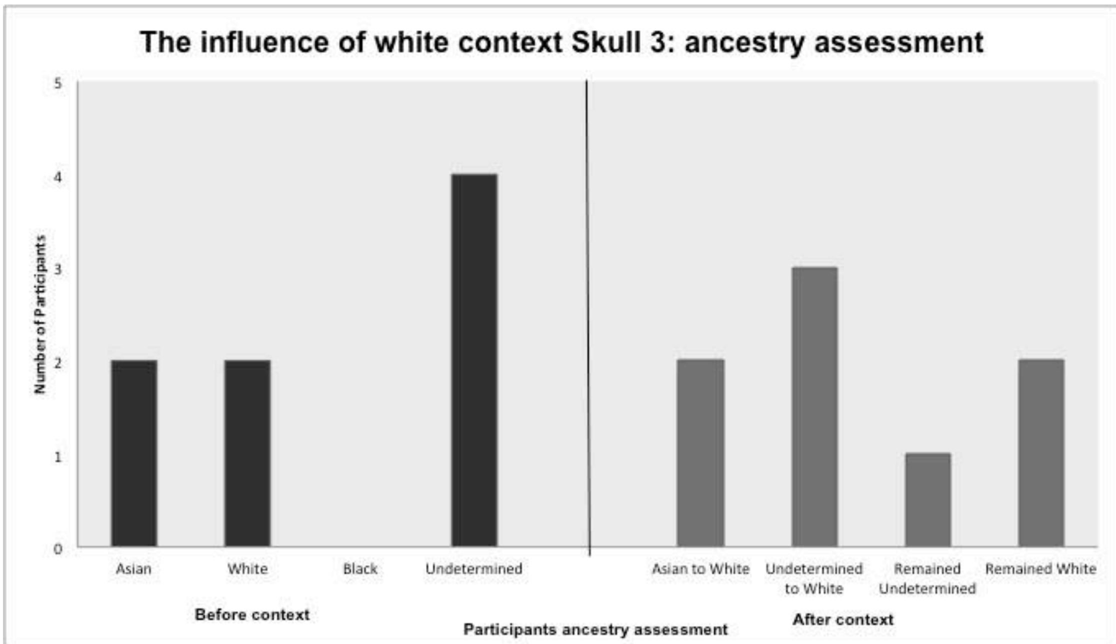

Figure 3.12 showing the interpretations of the skeletal remains before and after receiving White context, demonstrating a change in the interpretations (to align with the context) for five participants.

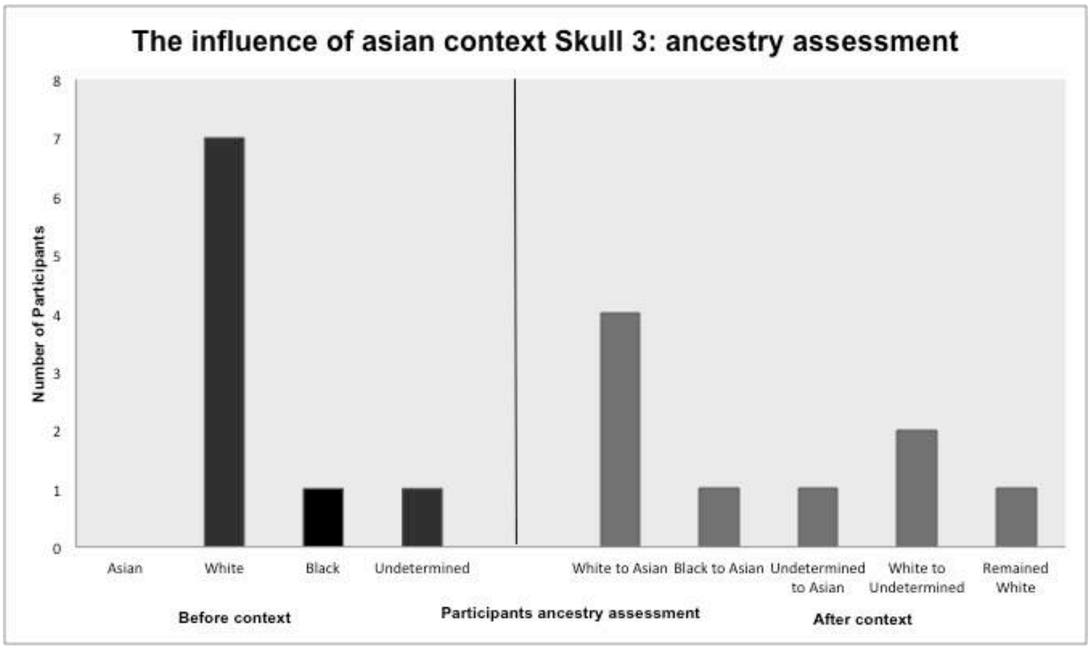

Figure 3.13 showing the interpretations of the skeletal remains before and after receiving Asian context, demonstrating a change in interpretations (to align with the context) for eight participants.

3.3.2.5. Descriptive statistics and Chi-square of os coxa 1

Sex assessment

A total amount of 17 participants were consistent with their previous decisions when comparing Baseline Control vs. Reliability Control. Out of the 17 consistent participants, 11 participants (65%) were affected by the contextual information, changing previous decisions, with 6 participants not being affected by the context (35%) (see Figure 3.14 and 3.15 for further details on the effect of female and male context). The results from the Chi-square test revealed a significant difference for participants given female context with a p-value <0.05. A significant difference was also found for participants given male context showing a p-value <0.05. See Table 3.6 for further details.

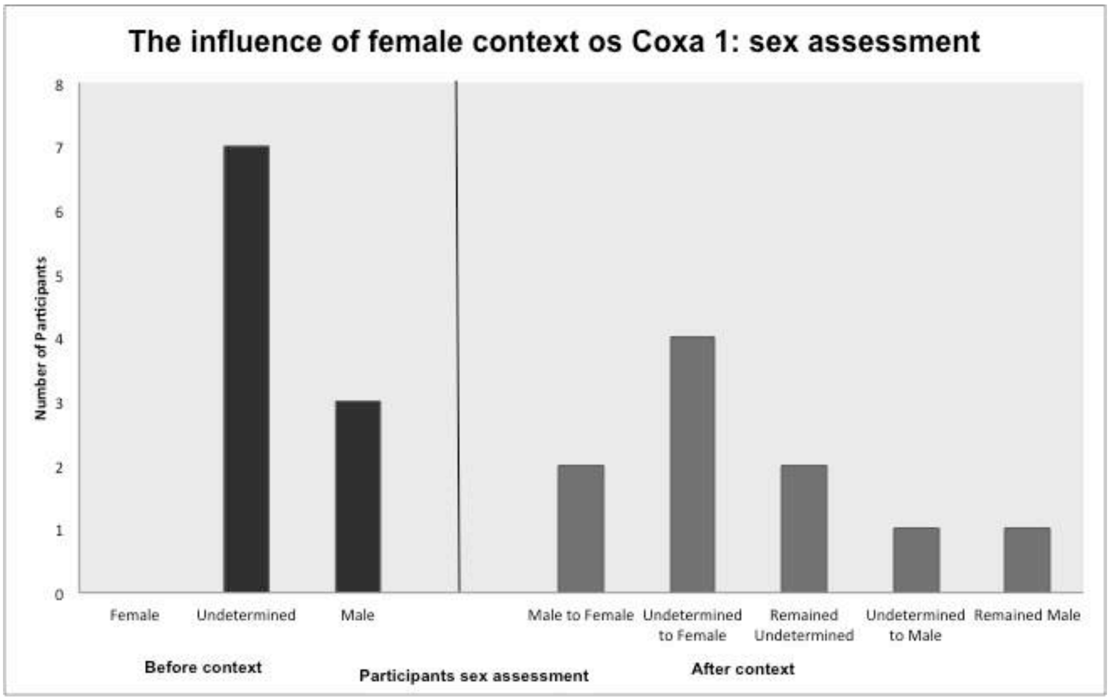

Figure 3.14 showing the interpretations of the skeletal remains before and after receiving female context, demonstrating a change in the interpretations (to align with the context) for six participants.

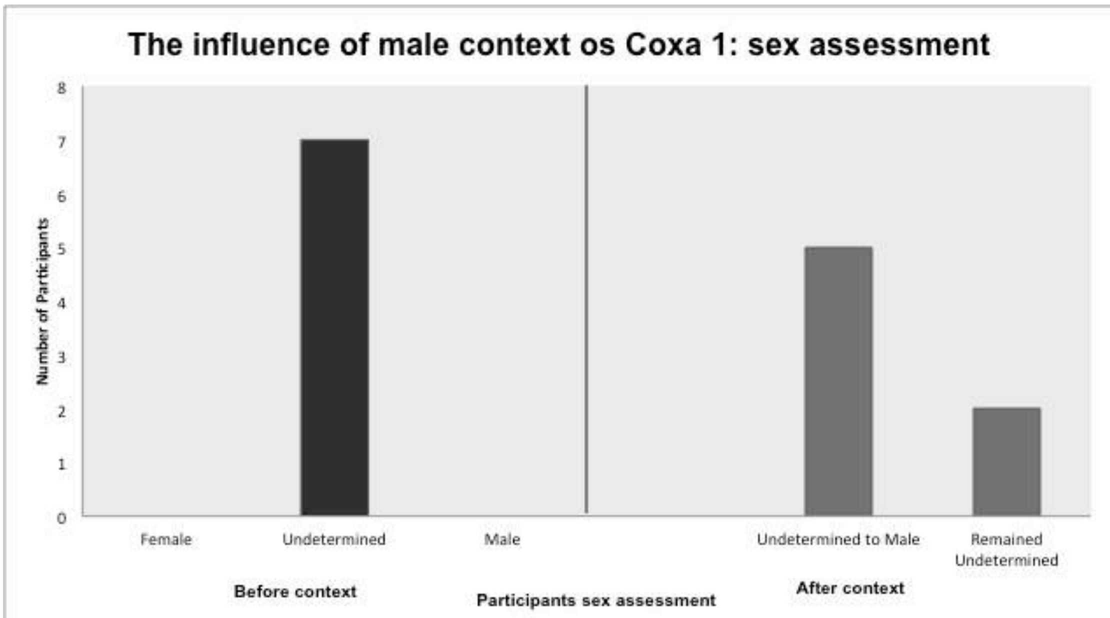

Figure 3.15 showing the interpretations of the skeletal remains before and after receiving male context, demonstrating a change in the interpretations (to align with the context) for five participants.

Age at death assessment

A total amount of 18 participants were consistent with their previous decisions when comparing Baseline Control vs. Reliability Control. Out of the 18 consistent participants, 9 participants (50%) were affected by the contextual information, changing previous decisions, with 9 participants not being affected by the context (50%) (see Figure 3.16 and 3.17 for further details on the effect of young and old context). The results from the Chi-square test revealed a significant difference for participants given context indicating a young individual with a p-value <0.05. A significant difference was not found for participants given context indicating an older individual, with a p-value >0.05. See Table 3.6 for further details.

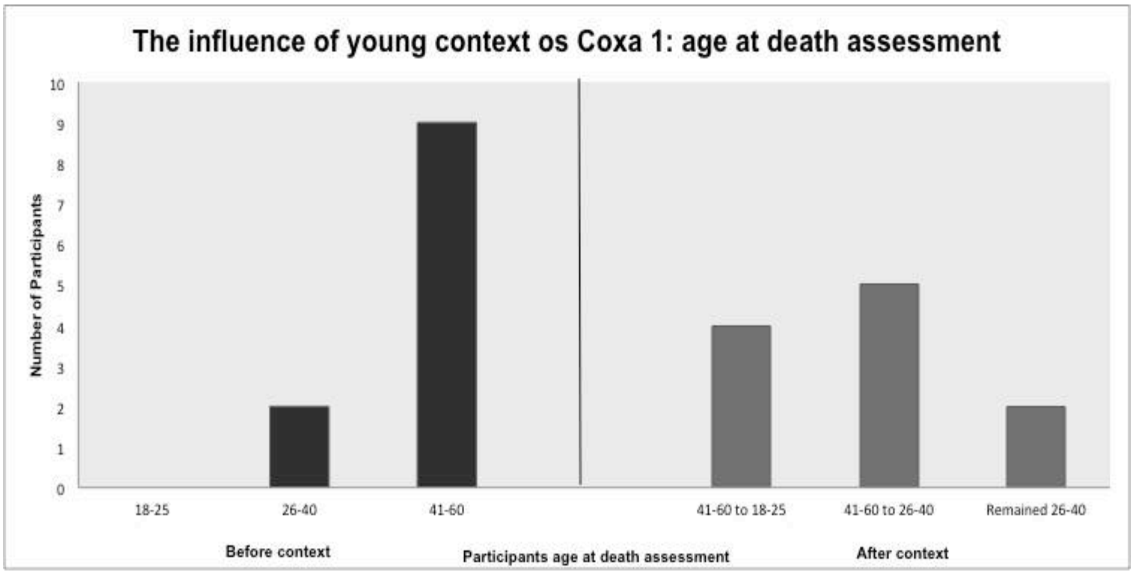

Figure 3.16 showing interpretations of the skeletal remains before and after receiving young context, demonstrating a change in the interpretations (to align with the context) for nine participants.

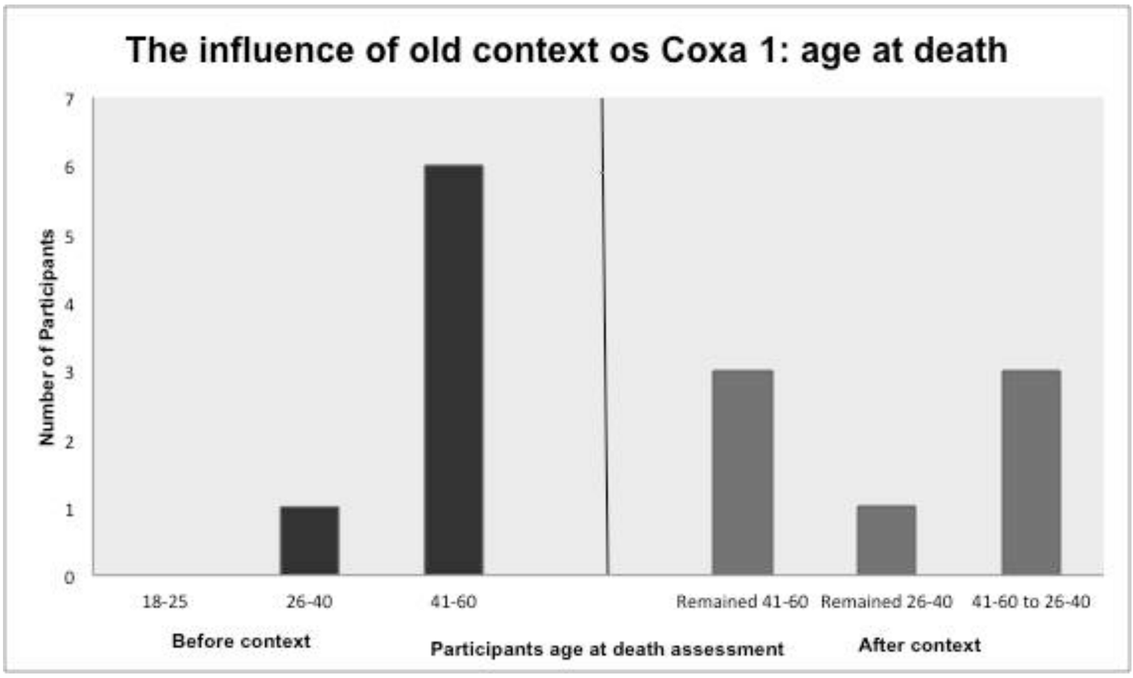

Figure 3.17 showing the interpretations of the skeletal remains before and after receiving context, demonstrating no change on the interpretations (to align with the context) for any of the participants.

3.3.2.6. Statistical Analyses and Chi-square of os coxa 2

Sex assessment

A total amount of 15 participants were consistent with their previous decisions when comparing Baseline Control vs. Reliability Control. Out of the 15 consistent participants, 9 participants (60%) were affected by the contextual information, changing previous decisions with 4 participants not being affected by the context (27%) and 2 participants (13%) being untestable (see Figure 3.18 and 3.19 for further details on the effect of female and male context). The 3 untestable participants were taken out for the Chi-square test when comparing Baseline/ Reliability Control vs. Context. The results from the Chi-square test revealed a significant difference for participants given female context with a p-value <0.005. A significant difference was not found for participants given male context indicating, with a p-value >0.05. See Table 3.6 for further details.

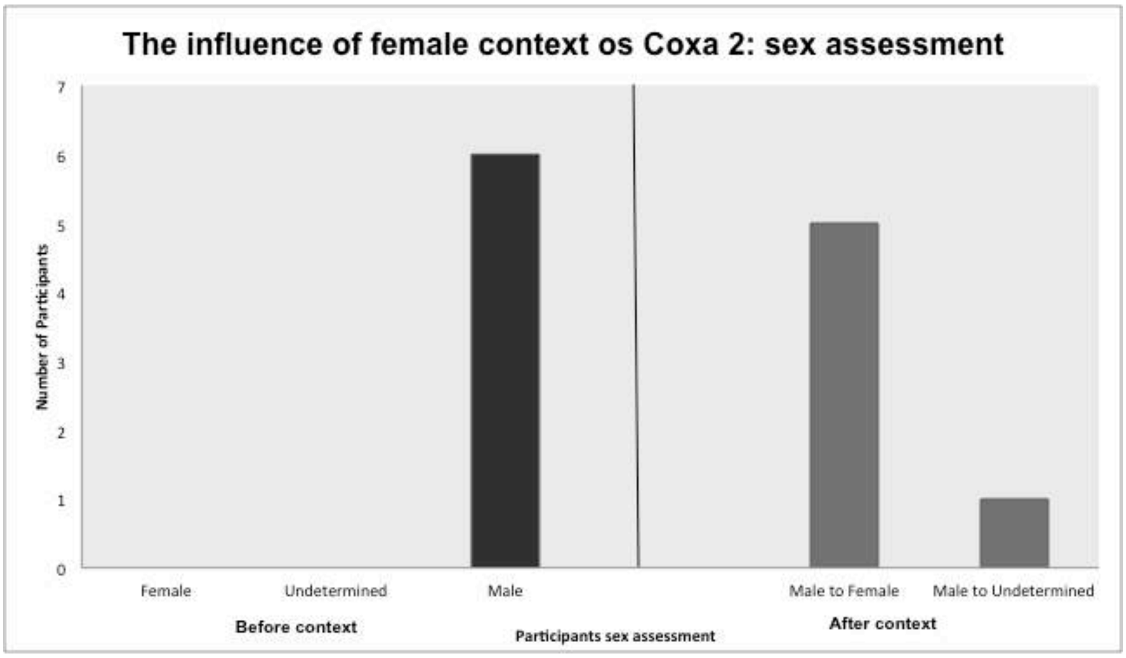

Figure 3.18 showing the interpretations of the skeletal remains before and after receiving female context, demonstrating a change in the interpretations (to align with the context) for all six participants.

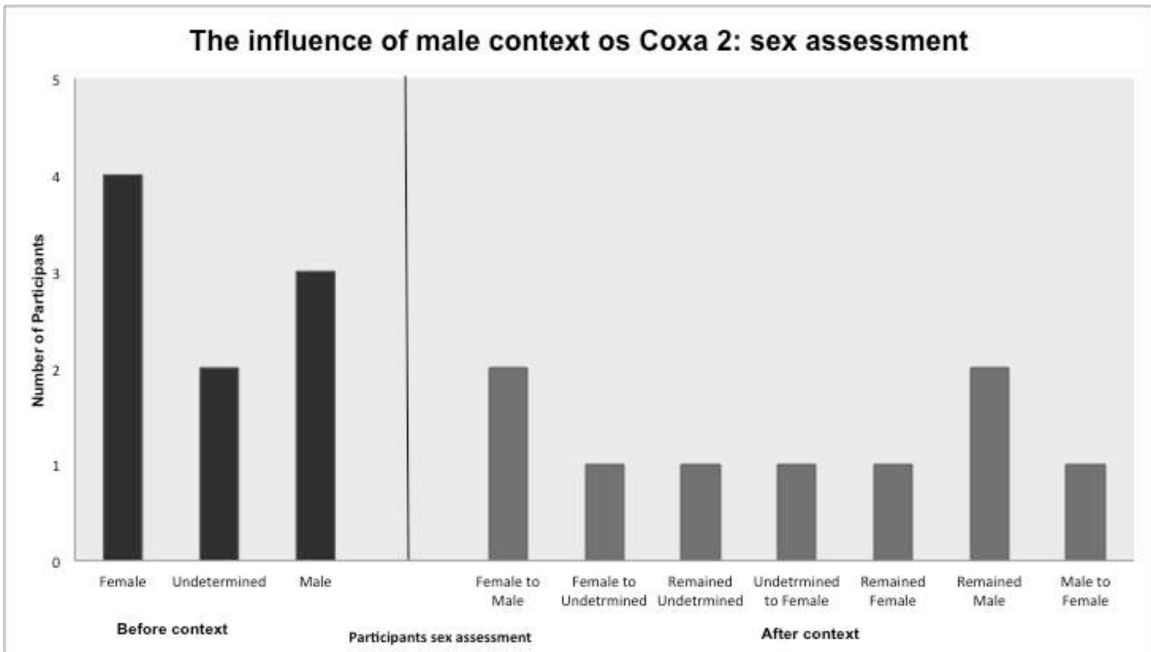

Figure 3.19 showing interpretations of the skeletal remains before and after receiving male context, demonstrating a change in the interpretations (to align with the context) for four participants.

Age at death assessment

A total amount of 18 participants were consistent with their previous decisions when comparing Baseline Control vs. Reliability Control. Out of the 18 consistent participants, 10 participants (55%) were affected by the contextual information, changing previous decisions, with 6 participants not being affected by the context (33%) and 2 participants (12%) being untestable (see Figure 3.20 and 3.21 for further details on the effect of young and old context). The 2 untestable participants were taken out for the Chi-square test when comparing Baseline/ Reliability Control vs. Context. The results from the Chi-square test revealed a significant difference for participants given a young context with a p-value <0.05. A significant difference was not found for participants given old context indicating, with a p-value >0.05. See Table 3.6 for further details.

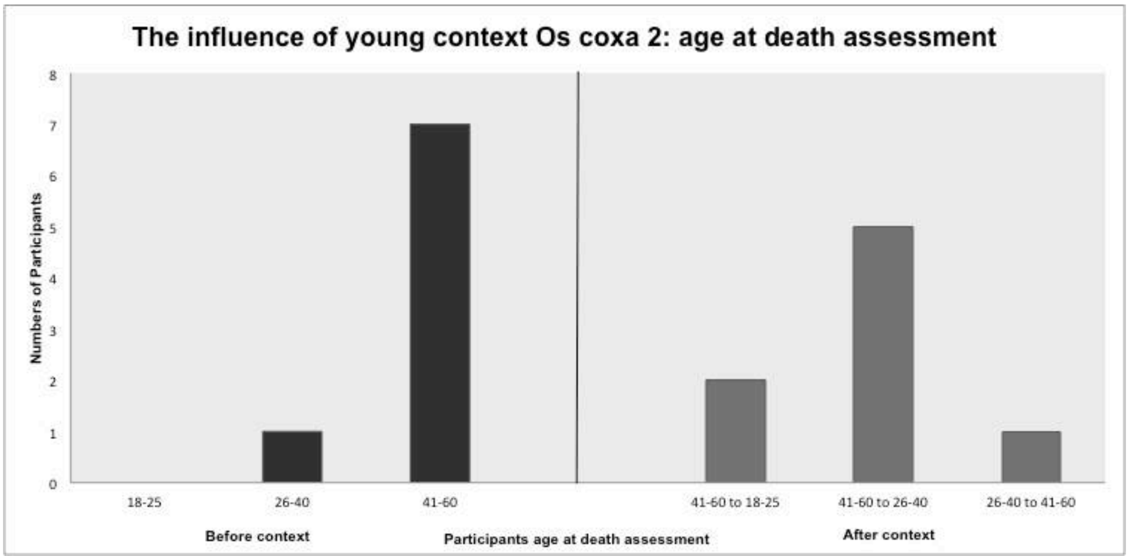

Figure 3.20 showing the interpretations of the skeletal remains before and after receiving young context, demonstrating a change in the interpretations (to align with the context) for seven participants.

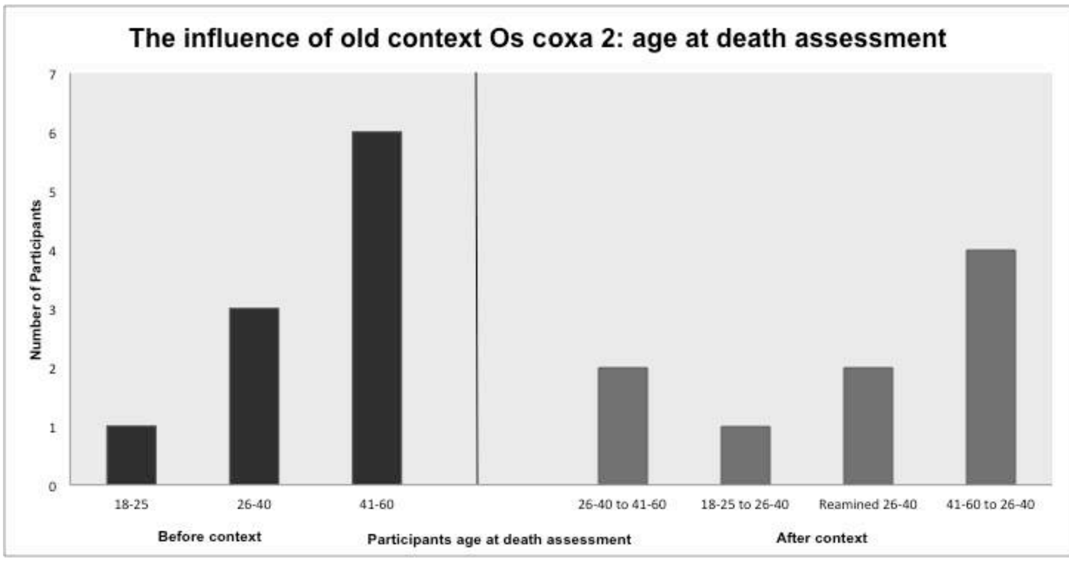

Figure 3.21 showing the interpretations of the skeletal remains before and after receiving old context, demonstrating a change in the interpretations (to align with the context) for three participants.

Chi square assessments Skull 1	Pearson's Chi-square	Asymp Sig.	Exact Sig.
Sex assessment			
Baseline Control vs. Reliability Control (n=44)	.472	.790	.846
Baseline Control vs. Female context (n=20)	7.200	.027	.048
Baseline Control vs. Male context (n=12)	3.133	.209	.351
Ancestry assessment			
Baseline Control vs. Reliability Control (n=44)	.730	.866	.886
Baseline Control vs. White context (n=10)	5.200	.158	.167
Baseline Control vs. context (n=16)	6.571	.037	.041
Chi-square assessments Skull 2	Pearson's Chi-square	Asymp Sig.	Exact Sig.
Sex assessment			
Baseline Control vs. Reliability Control (n=44)	2.849	.241	.270
Baseline Control vs. Female context (n=12)	5.467	.065	.080
Baseline Control vs. Male context (n=14)	10.500	.005	.005
Ancestry assessment			
Baseline Control vs. Reliability Control (n=44)	.598	.897	.925
Baseline Control vs. White context (n=28)	21.667	.000	.000
Chi-square assessments Skull 3	Pearson's Chi-square	Asymp Sig.	Exact Sig.
Sex assessment			
Baseline Control vs. Reliability Control (n=44)	3.826	.281	.270
Baseline Control vs. Female context (n=18)	10.500	.005	.002
Baseline Control vs. Male context (n=14)	3.943	.139	.143
Ancestry assessment			
Baseline Control vs. Reliability Control (n=44)	4.471	.215	.233
Baseline Control vs. White context (n=12)	8.800	.012	.026
Baseline Control vs. Asian context (n=20)	11.700	.008	.004

Table 3.5 showing Chi-square and Fisher's exact test results for all three skulls

Chi square assessments os coxa 1	Pearson's Chi-square	Asymp Sig.	Exact Sig.
Sex assessment			
Baseline Control vs. Reliability Control (n=44)	4.476	.215	.178
Baseline Control vs. Female context (n=20)	8.978	.011	.011
Baseline Control vs. Male context (n=14)	7.143	.008	.029
Age at death assessment			
Baseline Control vs. Reliability Control (n=44)	3.514	.173	.212
Baseline Control vs. Young context (n=22)	6.500	.039	.037
Baseline Control vs. Old context (n=14)	3.818	.051	.192
Chi-square assessments os coxa 2	Pearson's Chi-square	Asymp Sig.	Exact Sig.
Sex assessment			
Baseline Control vs. Reliability Control (n=44)	1.510	.470	.549
Baseline Control vs. Female context (n=12)	10.000	.007	.004
Baseline Control vs. Male context (n=14)	.476	.788	1.000
Age at death assessment			
Baseline Control vs. Reliability Control (n=44)	2.384	.304	.415
Baseline Control vs. Young context (n=16)	7.778	.020	.013
Baseline Control vs. Old context (n=16)	1.077	.299	1.000

Table 3.6 showing Chi-square test results for os coxa 1 and 2 skulls

3.3.3. *Comparing confidence level*

A one way repeated measure ANOVA was conducted to compare the confidence level before, during, and after resaving contextual information on participants being affected by the context and changing their initial decision-making. This was followed up by a paired sample t-test to make a post hoc comparison between conditions in order to see where the change might have taken place. An overall summary of the mean confidence value across all skeletal elements between all three groups is outlined in Table 3.7.

Table 3.7 An overall summary of the mean confidence value across all skeletal elements between all three groups

Skeletal element	Mean(%)	SD	N	Skeletal element	Mean(%)	SD	N
Skull 1 sex				*Skull 1 ancestry*			
Baseline-control	65.62	11.087	16	Baseline-control	60.33	20.569	15
Context	69.69	9.214	16	Context	68.67	11.095	15
Reliability-control	62.81	8.750	16	Reliability-control	60.33	12.459	15
Skull 2 Sex				*Skull 2 Ancestry*			
Baseline-control	71.56	14.913	16	Baseline-control	63.42	12.140	19
Context	68.75	17.272	16	Context	73.89	11.469	19
Reliability-control	68.75	15.111	16	Reliability-control	68.68	13.829	19
Skull 3 Sex				*Skull 3 Ancestry*			
Baseline-control	69.72	14.600	18	Baseline-control	65.88	13.019	17
Context	72.50	15.741	18	Context	71.18	17.187	17
Reliability-control	63.61	11.607	18	Reliability-control	65.59	13.565	17
Os coxa 1 sex				*Os coxa 1 age at death*			
Baseline-control	62.50	19.235	16	Baseline-control	61.67	17.150	18
Context	64.69	19.788	16	Context	65.83	13.089	18
Reliability-control	60.00	21.833	16	Reliability-control	65.83	15.554	18
Os coxa 2 Sex				*Os coxa 2 Age at death*			
Baseline-control	67.00	15.213	15	Baseline-control	56.78	15.201	17
Context	73.67	9.155	15	Context	67.35	9.206	17
Reliability-control	60.00	9.730	15	Reliability-control	56.76	15.303	17

3.3.3.1. *Confidence level Skull 1 sex assessment*

There was no significant effect on confidence level between the time before, during and after context, *Wilks' Lambda=0.760,F (2.11) =1.738, p=.221.*

3.3.3.2. *Confidence level skull 1 ancestry assessment*

There was a significant effect on confidence level between the time before, during, and after context, *Wilks' Lambda=.177,F (2.7) =16, p=.002.*

The result from the first paired sample t-test indicated that there was no significant difference between confidence level before (Mean=60,SD=22) and during

(Mean=68, SD=11.) context, *t;(8)=-1.325, p=.222*. Similar results were found for the second paired t-test before (Mean=60, SD=11) and after (Mean=61, SD=11) context; *t(8)=-.1741, p=.866*. However, as significant difference was established during (Mean=68, SD= 11) and after (Mean=61, SD=11) context; *t(8)= 5.965, p=.000*.

3.3.3.3. Confidence level Skull 2 sex assessment

There was no significant effect on confidence level between the time before, during and after context*, Wilks' Lambda=.667,F (2.8) =1.994, p=.297.* See table for further information.

3.3.3.4. Confidence level Skull 2 Ancestry assessment

There was a significant effect on confidence level between the time before, during and after context, *Wilks' Lambda=.548,F (2.10) =4.128, p=.049.*

The result from the first paired sample t-test indicated that there was a significant difference between confidence level before (Mean=62,SD=12,) and during (Mean=74, SD=8) context; *t(11)=-.3.014, p=.012*. For the second paired and third sample t test no significant difference was found before (Mean=62, SD=12) and after (Mean=71, SD=13) context; *t(11)=-1.789, p=.101* as well as during (Mean=74, SD= 8) and after (Mean=71, SD=13) *context; t(11)= .889, p=.393*.

3.3.3.5. Confidence level Skull 3 sex assessment

There was a significant effect on confidence level between the time before, during, and after context, *Wilks' Lambda=.377,F (2.9) =7.42, p=.012.*

The result from the first paired sample t-test indicated that there was no significant difference between confidence level before (Mean=68,SD=15) and during (Mean=72, SD=17.) *context, t;(10)=-.987, p=.347*. Similar results were found for the second paired t test before (Mean=68, SD=15) and after (Mean=60, SD=12) context; *t(10)=2.136, p=.058.* However, as significant difference was established during (Mean=72, SD= 17) and after (Mean=60, SD=12) *context; t(10)= 3.938, p=.003.*

3.3.3.6. Confidence level Skull 3 ancestry assessment

There was no significant effect on confidence level between the time before, during and after context, *Wilks' Lambda=.859,F (2.11) =.899, p=.167*. See table for further details.

3.3.3.7. Confidence level os coxa 1 sex assessment

There was no significant effect on confidence level between the time before, during and after context, *Wilks' Lambda=.938,F (2.9) =.297, p=.084*. See table for further details.

3.3.3.8. Confidence level os coxa 1 age at death assessment

There was no significant effect on confidence level between the time before, during and after context, *Wilks' Lambda=.635,F (2.8) =.298, p=.163.* See table for further details.

3.3.3.9. Confidence level os coxa 2 sex assessment

There was no significant effect on confidence level between the time before, during and after context, *Wilks' Lambda=.689,F (2.7) =1.578, p=.233.*

3.3.3.10. Confidence level os coxa 2 age at death assessment

There was no significant effect on confidence level between the time before, during and after context, Wilks' Lambda=.650,F (2.6) =1.346, p=.341. See table for further details.

3.4. Discussion

3.4.1. Experiment 1 (Pilot study)

The initial findings of the Pilot study indicate that additional grave context descriptions and artifacts as well as osteological reports, can have an influence on the judgments of participants resulting in a change to their previous decisions. The results showed that when context was given, (prior to the assessments), participants changed their decision-making (76.8.%), confirming with the contextual information.

Sex assessment

Furthermore, the decision-making change of the participants was notable across all skeletal elements within every step of the biological profile. For example in sex assessment of skull 1 only two participants (13.3%) made a consistent analysis regardless of the context provided, thirteen participants (86.7%) provided an answer that confirmed the context given and changed their previous decisions. Similar results were found in sex assessment of skull 2 and skull 3 demonstrating a possible vulnerability to contextual influences within the visual methods used in sex estimation on the skull. In addition, it also demonstrated that both types of contextual information affected the participants (male and female) showing not only the level of ambiguity in the morphological features but also that participants tended to rely on any type of context when making interpretations, arguably looking for features that would be more in accordance with the context given. The sex assessments on the os coxa also demonstrated a change within participant's decision-making when context was given. In general, participants were more consistent within sex assessment on the os coxa in comparison with the skulls (which could be due to the os coxa being 'less' ambiguous in morphological traits compared to the skull).

Ancestry

Similarly, in ancestry estimations majority of participants had a tendency to include the context in their interpretations of the skeletal remains, resulting in change on pervious ancestral morphological features For example, on skull number 2, all ten participants changed their previous decisions to agree with the context provided. Traditional approaches to ancestry assessments in forensic anthropology have been scrutinized to be subjective by nature (Hefner et al. 2012). This pilot study suggests that assessing ancestry from discreet traits is not an easy undertaking, with difficulties establishing a single known trait to be exclusively found in only one population. For example, only seven participants in total (across all ancestry assessments on all three skulls) assessed the ancestry to be ‘undetermined’ on previous decisions. The majority of participants either estimated the skulls to be of Caucasian, Asian or Black descent. For example on skull 1, four participants estimated, on previous decisions the skull to be of Caucasian descent, with another four participants assessing the same skull to be of Asian descent. Participants estimating skull number 1 to be of Caucasian descent was given a context indicating Asian ancestry and vice versa for participants estimating it to be of Asian ancestry, who were given a context indicating a Caucasian individual. All eight participants changed their prior interpretations of the skeletal remains to confirm with the Asian or Caucasian context given. This was also demonstrated throughout the ancestry assessments of skull 2 and 3. The result not only preliminary suggests an inconsistency within ancestry assessments but also an indication of a strong contextual influence in the decision-making process.

Age at death

In assessment of age at death, a change of participants previous decisions were notable when contextual information was given. For example on os coxa 12 participants (70.6%) changed their previous decisions confirming with the contextual information, with similar results shown in age at death assessment on os coxa 2. The majority of participants either estimated the individual to be younger than in their previous decision or older depending on the context given. Age at death assessments in forensic anthropology have been scrutinised for having a significant discrepancy between observers and issues have been raised with regards to lack of appropriate techniques and statistics associated with age at death within a forensic context (Lottering et al. 2013; Hoppa & Vaupel 2002). This study suggests that visual assessments of age at death are also vulnerable to contextual information potentially causing an observer effect within participant’s assessments.

Limitations with Experimental study 1

Although the pilot study indicated that context could play a role in the interpretation process of what participants may observe, the limitations within the pilot study did not allow for further details on how much of the decision-making change was solely based on the contextual information, and not on participants being generally inconsistent, due to lack of experience, as studies within forensic anthropology has shown that experience do play a significant role in the application of anthropological methods. Even though the majority of participants were undetermined on their initial decisions, (which would be expected and arguably the right answer due to the ambiguity of the skeletal remains) yet, there was still some inconsistency amongst participants answer. By adding a reliability control phase it was possible to establish if participants are generally consistent in their analysis regardless of contextual influences, and then by adding context further look into the effect of contextual effects in the assessment of skeletal remains.

3.4.2. Experiment 2 (Control/follow up study)

In a manner akin to the pilot study and previous research regarding the effect of context on the interpretation of forensic evidence (Dror et al. 2006), the findings of this study indicate that additional grave context descriptions and artifacts as well as osteological reports, can have an influence on the judgments of participants resulting in a change to their previous decisions. The results showed a high consistency (77%) in the answers from participants when no context was provided (with a majority of the decisions being 'undetermined'). However, when context was given, prior to their assessment, the consistent participants changed their interpretation of the skeletal remains (63%), as a result of the contextual information. Although Experimental 1 showed a higher percentage in the effect of context on the interpretation of the skeletal remains (76.8%), Experimental study 2 still highlights the susceptibility of visual assessments to affects of contextual information, when making interpretations on ambiguous skeletal remains.

Contextual influences and sex assessments

Most importantly, the data demonstrate that even relatively 'weak' context can influence the decision-making outcome. Previous studies in forensic anthropology showed that contextual biases were present when a strong context was provided (such as DNA results, see Nakhaeizadeh, Dror, et al. 2014). This present study however did not include such a strong contextual influence, but nevertheless showed the vulnerability to contextual information when establishing a biological profile when decisions are being made on ambiguous skeletal remains. For example, the results from the sex assessment on skull 1 demonstrated that 81% of the sixteen consistent participants were affected by the contextual information when assessing the sex of the skull, with significant changes occurring on the previous decision for the ten participants receiving contextual information indicative of a female. The contextual information for the ten participants receiving a female indicative context included a grave setting description of a whalebone, an artifact commonly used for corset stay in medieval periods. It is plausible that the preconceived idea of participants that whalebone artifacts may be more associated with female graves might have affected their interpretation of the skeletal remains, resulting in a change on previous decisions and the interpretation of the skeleton. Similarly, for the six participants on skull 1 sex assessment receiving male indicative contextual information (such as a buckle belt) resulted in five participants changing their initial decisions to confirm with the male context with only one participant going against the context. Although the male context for skull 1 sex assessment did not show a significant difference (which could have been due to the small sample size of participants receiving male contextual information), the majority of participants in this group still changed their previous decisions even when provided with a less strong context (in comparison to DNA). This was also notable for sex assessment across skull 2 (62%), skull 3 (62%), os coxa 1 (65%), and os coxa 2 (60%) with some of the results for the os coxa demonstrating both female and male context having a statistically significant influence on participants sex estimation of the skeletal remains.

Contextual influences and ancestry assessments

Similar results were found in ancestry estimation. For example, ancestry estimation for skull 3 showed 76% of the seventeen consistent participants being affected by the contextual information, when making ancestry estimations with a significant effect of context on participants receiving contextual information indicative of White ancestry as well as Asian ancestry. The results indicate that both types of contextual information led to a change in the interpretation of the ancestral morphological traits as a result of the context. Comparable results were shown for ancestry estimation on skull 1 (56%) and skull 2 (63%) demonstrating the majority of participants changing previous ancestry estimations when context was provided.

Contextual influences and age at death assessments

Equally, for age at death assessment on os coxa 1, 50% of the eighteen consistent participants changed previous decisions with a significant difference in answers for participants receiving contextual information indicative of a younger individual. This change was also notable for os coxa 2 with 55% of the seventeen consistent participants changing their previous decisions, again only with a significant difference in participants receiving contextual information indicative of a young individual. Age at death assessments in forensic anthropology have a wide age range and sometimes an overlap between phases for the methods used on the os coxa. It is plausible that this might have resulted in participants not needing to change previous age at death decisions when given context of an older indivdiual as it might have fallen within the same age range. Giving participants more narrow age ranges in the answering sheets to choose from would have been of value for this study.

Decision-making process

Even though the majority of participants tended to provide a conclusion on the skeletal remains that confirmed the context given in phase 2, thereby coming to a different conclusion in comparison to their context free decisions across all five skeletal elements, it is important to consider the initial decision-making process. For example, participants who made an initial definite decision (e.g. male or female) and subsequently changed their decision to fit the context given (e.g. male to female), arguably made a greater change in their interpretations, in comparison to participants who initially said an exhibit was undetermined in their initial decision, and then changed their decision to a classification of male or female. However, with the exception of age at death, many of the participants did not make a definite decision with some assessing the exhibits to be undetermined in their initial decision-making especially in sex assessment. This was not surprising as the skeletal remains chosen for this experiment were set out to be of ambiguous nature. By including the option ‘undetermined’, participants were given a wider choice than a binary male or female forced-choice decision. Therefore it is perhaps not surprising that when assessing the ambiguous skeleton with the mixed traits provided, the ‘undetermined’ option was preferred to a definite male or female classification. This could also explain the high percentage of consistency amongst participants when context was not given. Additionally, when indicative context was given regarding an ambiguous skeletal element, it was not unexpected that participants unconsciously relied on the contextual information to a greater degree when making interpretations, arguably focusing on traits that would confirm with the context given.

It is also important to highlight that the consistency in this case was measured on the basis of the overall answer from all the methods available and combined for each element, rather than on one technique or the single traits scored for each method. This is important, as previous validation and classifications studies within methods used in forensic anthropology have generally shown these methods to be reliable, with high classification accuracy, specifically for sex estimation on the pelvis (Spradley & Jantz 2011). The aim of both Experiments 1 and 2 was not to conduct a validation and classification study of non-metric methods used in forensic anthropology, but rather to investigate further into the role context may play in the visual assessment of skeletal remains when making decisions on ambiguous skeletal remains. The reliability-control added to Experimental study 2 was designed not to assess reliability in 'judgment accuracy' (as that was not possible due to lack of a known sample size) but to see if participants would be consistent in their interpretations regardless of context.

As mentioned in Chapter 2 section 2.3 a recent study in forensic anthropology focusing on the innominate has however shown that there is a tendency to change the scaling of single traits on the innominate that have been assessed in isolation, to fit the overall decision reached. The study indicated a confirmation bias where the overall appearance of skeletal elements could affect previous scoring of traits conducted in isolation (Klales & Lesciotto 2016). Although Experiment 1 and Experiment 2 did not focus on the single traits, it is plausible that participants tend to rely more upon the traits that were more in accordance with the indicative contextual information. A valuable recommendation study would be to further explore each visual method separately in order to establish possible cognitive biasing influences inherent in each individual method within a known sample.

The confidence level for each assessment allowed for a further understanding of how ascertain participants were when making decisions during the different times of the experimental study, (before, during and after receiving context). The result shows that in general participants were most confident when having contextual information available prior to their assessment. With an exception of Skull 1 sex assessment, the mean confidence value across all skeletal elements was generally higher during the context round. This might not be surprising as many of the participants were undetermined in their initial decision-making as such, when given context, participants might have felt more confident making a final decision in accordance with the contextual information. It is important to highlight that the significant difference of the confidence level was only detected within skull 1 2 and 3 and only for some of the analysis, yet it still showed that the significant difference was almost always detected between the stages of during and *after* context where the confidence level of the majority of participants lowered in the last stage of the experiment. This is interesting, as generally it is believed that confidence increases over time, as more experience and practice is 'gained'. However, in this experiment arguably it appears to have decreased in the decision-making process. As mentioned previously, this could be due to participants feeling more confidant making a definite decision on an ambiguous skeletal remain when context was provided and felt less confidant in their decision-making process in absent of context.

Limitations

Similar to Experimental study 1, the current study was limited in terms of the sample size used due to participant availability, however, 22 participants was considered to offer valuable insights in this preliminary study. The chi-square results across sex, ancestry and age at death estimations showed that the difference in decisions with context (in phase 2) was more notable within sex and ancestry estimation, and in some cases within a certain indicative context. This could be due to smaller effect sizes within some groups, hence the sample size not being sufficient to reflect the effects statistically. The difference in decisions with context being more notable within certain skeletal elements could also be a result of the level of 'ambiguity' present (or absent) for each skeletal element on certain traits, as well as the type of contextual information provided. Likewise, providing participants with a time limit on the assessment of the skeletal remains (due to the workshop design), as well as setting up the task as a practical mini-test as part of the workshop, may have affected the interpretation of the skeletal remains made by the participants.

Moreover, it is also important to acknowledge that participants in both Experiment 1 and Experiment 2 were not working experts within the field of forensic anthropology, A valuable comparable study would be to investigate if similar effects were to be identified amongst working forensic anthropologist professionals. Previous work in forensic anthropology (as well as in other expert domains) have however showed that context effects do indeed affect working experts (Nakhaeizadeh, I. . Dror, et al. 2014), especially when they do not know they are being tested and when they believe the contextual information (Dror 2011).

3.5. Conclusion

The results presented in this chapter demonstrate that context influenced the decision-making of participants when evaluating previously assessed skeletal remains, revealing a potential for preconceptions to influence the interpretation of the skeletal remains. Similar to the concerns raised by Walker (1995), Weiss (1972), and Effros (2000) with regards to expectation bias in skeletal assessments, which may affect the interpretation of the skeletal remains, the results from both experimental studies 1 and 2 indicate that gravesite artifacts as well as osteological reports could create a preconceived idea of a certain sex, ancestry and age at death. This might also possibly indicate that participants rely more on the contextual information than one might account for. Indeed it raises the question of the subjectivity in the methods applied in forensic anthropology when decisions are being made on ambiguous skeletal remains, showing in both studies that the decision-making outcome was arguably more based on the contextual information rather than the scientific interpretation.

Chapter 4. Initial exposure to extraneous factors at the crime scene and subsequent bias in the processing of skeletal remains.

4.1. Introduction

While the forensic community is progressively accepting the importance and relevance of human cognition and decision-making, the debate on how to control and minimise unconscious contextual biases is still an open issue (as outlined in section 2.2). The issue of how to increase objectivity in criminal investigation at an early stage has been intensified, (de Gruijter et al. 2016; van den Eeden et al. 2016), with a growing number of documentaries drawing public attention and highlighting the consequences of these potential biases, affecting evidence collection, interpretation, and presentation in a court of law (Innocence project 2017; FBI & Federal Bureau of Investigation 2015; W. C. Thompson 2009; Ricciardi & Demos 2015). Indeed, the criticism and discussions in the literature have mainly focused on the biasing effect of domain irrelevant information influencing the decision-making of experts, with proposed solutions pushing for the need to minimise task irrelevant contexts (Dror et al. 2015). Many of the recommended solutions are targeting different disciplines within forensic science (Kerkhoff et al. 2015; Krane et al. 2008; Archer & Wallman 2016) at different stages in the forensic science process (Thompson 2011; Edmond et al. 2014; van den Eeden et al. 2016).

However, what is considered as relevant and irrelevant information when making forensic interpretations is not always an easy undertaking. Furthermore, it has been argued that there is benefit in exposing the scientist to contextual information, and that mitigating bias by detaching the science from the criminal process is in fact a disadvantage (Champod 2014). Others suggest that such exposure is good for motivating forensic examiners and for their 'personal satisfaction' (Butt 2013). Further concerns have been raised with regards to the fact that research into subjective decision-making might detract from focusing on increasing the objectivity with which forensic evidence can be interpreted, (Champod 2014) for example through an improved understanding of the dynamics of forensic trace materials (Morgan et al. 2014) . Nevertheless, there are many crucial decisions being made throughout the progression of evidence from crime scene to court (Dror 2015). The empirical evidence base that underpins how one makes decisions, what influences those decisions, and how to enhance decision-making outcomes, is still not fully appreciated in all forensic domains at all stages of a criminal investigation.

In forensic anthropology very little is known about how early exposure of context might affect the subsequent assessment of skeletal remains. Like other forensic domains, exposure to environmental and contextual influences varies, depending on the particular case, organizational practice and procedures, and the nature of the forensic domain. In some cases forensic anthropologists may be called to the crime scene in order to provide on-site identification of skeletal remains (Cheetham & Hanson 2009), be part of the revision of search strategies (Haglund 2001), as well as helping to preserve, excavate, and document the skeleton in situ (Cheetham & Hanson 2009). This is of importance as the expertise and knowledge of the forensic anthropologists on site can significantly aid in the outcome of a death investigation. However, this could also potentially create an early exposure to a potentially significant amount of context that may in some cases, be considered as 'task irrelevant' and have the potential to cause bias in interpretation at a later stage. Some have argued that there might be a potential for expectation bias in the interpretation of skeletal remains when exposed to context, specially when making assessments on ambiguous skeletons (Nakhaeizadeh, Hanson, et al. 2014).

This chapter will therefore focus on the potential cascading effects of initial exposure to extraneous context at a crime scene upon subsequent judgment and decision-making. More specifically, this was done in order to addressed whether clothing associated with skeletal excavations at the crime scene could influence and impact the evaluations and judgments of participants. This was to be done in order to examine whether early exposure to such contexts would cascade and affect the subsequent assessment of the skeletal remains in the laboratory.

4.2. Methodology

4.2.1. Experimental design

Participants in this study investigated a mock crime scene, focusing on forensic archaeological techniques and the excavation of clandestine burials, followed by a 'forensic anthropological' assessment on the skeletal remains. The experiment was designed in order to research whether initial exposure to extraneous grave context had an influence on the primary working hypotheses, and thereafter the assessment of the skeletal remains, focusing on the estimation of sex. The experiment was carried out in three phases, with a three-month interval between each phase:

- Phase one: the preparation and the burial of the skeletal remains,
- Phase two: the excavation and assessment of the skeletal remains,
- Phase three: a control study, in which participants assessed the skeletal remains blindly

This was repeated over a period of two years to increase the number of participants.

4.2.2. Material

Four identical disarticulated casts of the human skeleton representative of white males were used in this study. The same four casts were used to replicate the study the following year. Casts of human skeletal remains are regularly used in medical schools, forensic anthropology, and osteology courses as teaching materials in lieu of real skeletons. Therefore, the morphological features on the casts used in this study possessed very distinctive male characteristics, with very few ambiguous features present.

The casts were dressed in clothes prior to burial, with two of the male skeletons dressed in female clothing, and two dressed in gender neutral garments, i.e. perceived as either male or female (see Figure 4.1 for an example of skeletal casts dressed in female clothing (a & b) and Figure 4.2 for gender neutral garments (c & d). This was in order to see if 'extraneous' clothing associated with skeletal excavations (for example female clothing on a male skeleton) could have an impact upon the early hypothesis, which could later cascade and impact interpretation and decision-making about the sex assessment at the later stage of the analysis. Furthermore, the use of a very strong context such as female clothes as opposed to a more ambiguous context (gender neutral clothing), allowed for a comparison within different types of contextual influences, as studies have repeatedly shown that people tend to hold on to their initial beliefs even if contradictory evidence is presented (e.g. Anderson and Kellam 1992). In addition, the skeletal remains were all buried with 'neutral' artifacts associated with each burial. Similar items were included in each of the graves such as contact lenses, mobile phones, SD cards, train tickets, cigarette stubs and coins (see Table 4.1).

Table 4.1 showing the neutral artifacts associated with each burial

	Grave A	Grave B	Grave C	Grave D
Artifacts	1 Necklace	1 Earing	1 Earing	1 Necklace
	1 Eye contact lens	1 Eye contact lens	1 Eye contact lens	1 Eye contact lens
	5 Coins	5 Coins	5 Coins	5 Coins
	1 Black empty wallet	1 Train ticket	1 Train ticket	1 Black cell phone
	1 Sd card	1 Cigarette stump	1 Cigarette stump	1 Sd card
	Clump of hair	Clump of hair	Clump of hair	Clump of hair

Figure 4.1 showing the burials of the skeletal casts dressed in female clothing

Figure 4.2 showing the burials of skeletal casts dressed in gender-neutral garments

4.2.3. Participants

A total number of 38 MSc students participated in this study, all with a bachelors degree and background in bioarcheology/biological and physical anthropology or osteology, with training and experience in excavations and the use of osteological techniques on skeletal remains. In order to minimise any potential influence on the decision-making process, participants in this study were not informed of the true nature of the experiment. The exercise was therefore included as part of a forensic archaeology module, in which the final examination and assessment of the module included taking part in a three-day mock crime scene excavation. The course was run over a period of eleven weeks with the course culminating in a simulation exercise of a serious crime investigation. Incorporating this study in the module also ensured that students took the exercise seriously and were motivated to keep errors to a minimum, as they were being assessed on their performance. The students on the forensic archaeology course, with previous background knowledge of forensic anthropological/osteological assessments were further asked to take part in the subsequent forensic anthropological analysis post excavation. This was set up in a mock mortuary facility. The participants were told that this was a mock mortuary exercise following the excavation, and to focus solely on the assessment of the skeletal remains.

4.2.4. *Procedure*

Phase one: Preparation of the burials and the mock crime scene

Four clandestine burials were created with each grave having an approximate diameter of 120x80cm, with a depth of roughly 20cm. Each grave included one fully clothed male skeletal cast with associated grave artifacts. All graves were identical in shape and with similar grave artifacts, the only difference being the clothing associated with the skeletal remains. Each individual skeleton was blindfolded as well as being bound by the feet and wrists, with imitation blood spattered on parts of the clothes. Each grave was filled and covered in order for the students to be able to locate and excavate the burials 2 months after the burials were created (see Figure 4.3 for illustration of the preparation). This procedure was replicated for the following year.

Figure 4.3 showing the preparation and completion of the clandestine burials and the mock crime scene

Phase two: Excavation and the assessment of the skeletal remains

The excavation of the skeletal remains took place over three days, with participants asked to locate the potential clandestine burials and excavate the graves accordingly. In order to make the exercise as close as possible to a real crime scene excavation, all participants had to follow protocols, chain of custody and standards accordingly, with logs and entrance points being observed at the scene of crime (see Figure 4.3 and 4.4). Participants were randomly assigned to groups of four/five with each group excavating one burial. Participants were asked to log, document and collect all evidence accordingly based on the training received from the forensic archaeology course. The skeletal remains were recovered and put in body bags and transported back to the mock mortuary.

Figure 4.4 showing the students finishing excavating one of the clandestine burials

Three mortuary stations were created in order to have more than one participant conducting the analysis at the same time Participants were first asked to document, remove and bag the clothing, and thereafter wash the skeletal remains. Participants were then asked to lay out the skeletal remains in anatomical order, and thereafter conduct a biological profile following the 'forensic anthropological report sheet'. The answering sheet report included most traditional common metric and non-metric methods used in forensic anthropological textbooks for sex, ancestry and age at death estimations. For the purpose of this study, the relevant results pertaining to the sex assessment were used (see Appendix C).

Participants were asked to follow and complete the report starting with visual assessments followed by metric analysis. In addition, participants were also asked to write any visible signs of pathology and trauma. At the end of the report, participants were asked to provide a short non-technical summary of their analysis on the skeletal remains. Additionally, in order to understand the decision-making process further, participants were also asked to provide a confidence level for each assessment and final evaluation of the skeletal remains. Participants were given access to reference materials and casts for the most common methods used in forensic anthropology for sex, ancestry and age at death estimations, as well as callipers and measurement boards for metric analysis. The time frame to conduct the analysis in the mortuary was approximately 45min-1hour.

Participants were told in advance that the mortuary exercise (along with other external exercises) was not assessed, as it was merely additional time to practice mortuary procedures in a forensic crime scene investigation. Participants were specifically told to not include any of interpretation of the skeletal remains in their final assessed reports, as it was not part of their forensic archaeological assessment.

Phase three: a control study

A control group was created and phase three of the study was run several months after the forensic archaeology module ended, with participants assessing the same male skeletal casts used in the previous exercise but in this phase, without any contextual influences. The skeletal remains were laid out (without any clothing or artifacts) in a lab facility, and participants with relevant background knowledge in forensic anthropological/osteological techniques were asked to establish a biological profile. Participants in this group had not previously taken part in the mock crime scene forensic archaeology exercise. This allowed for a comparison of answers between participants exposed to contextual influences compared to participants conducting the analysis in isolation. Participants were asked to fill in the same report sheet created for previous participants, and to conduct a full biological profile following the report, providing a non-technical summary of their findings, together with their confidence level. Participants in the control group were provided with the same access to the same reference materials for sex, ancestry and age at death estimations as previous participants.

After the completion of the assessments, participants in the control group were given a short summary of each burial. The short summary included information with regards to the skeletal remains being used in the mock crime scene exercise, as well as information with regards to the location of the skeletal remains, grave artifacts, and clothing associated with the skeletal remains, including both the female and the ambiguous clothing contexts. Participants were asked to fill in a short questionnaire, elaborating on whether their answers in respect of the additional information would change their previous decisions on the assessment of the skeletal remains, and their confidence in that decision-making process (See Appendix C). This was in order to assess whether the initial judgment of a participant would be affected by the additional information, or if participants would confirm their initial analysis without being influenced by the additional context provided.

4.2.5. Analysis

The data were recorded and analysed using descriptive statistics and SPSS for significant tests. In order to examine whether there was a significant difference between the groups as a function of the extraneous context, a series of Chi-square tests was carried out. Due to the small sample size a Fisher exact was also reported. Similar to Chapter 3, the variables under study in this experiment were each categorical and a Chi-square and Fisher's exact test was used to see if there would be a significant relationship between the sex assessments of the participants between the different groups depending on the two different context (female clothing vs. gender neutral clothing). Independent and dependent t-tests were used to compare the confidence level of participants between the different groups as well as within the control group.

4.3. Results

4.3.1. Decisions of all three groups

A total of 38 participants took part of the study with 11 participants in the female context group (Group 1), 12 in the ambiguous context group (Group 2) and 15 in the control group. Figure 4.3 shows the distribution of the decisions on the skeletal remains for all three groups sex assessment with 'male?' and 'female?' being the representative terms that indicate 'possibly' male and 'possibly' female in anthropology.

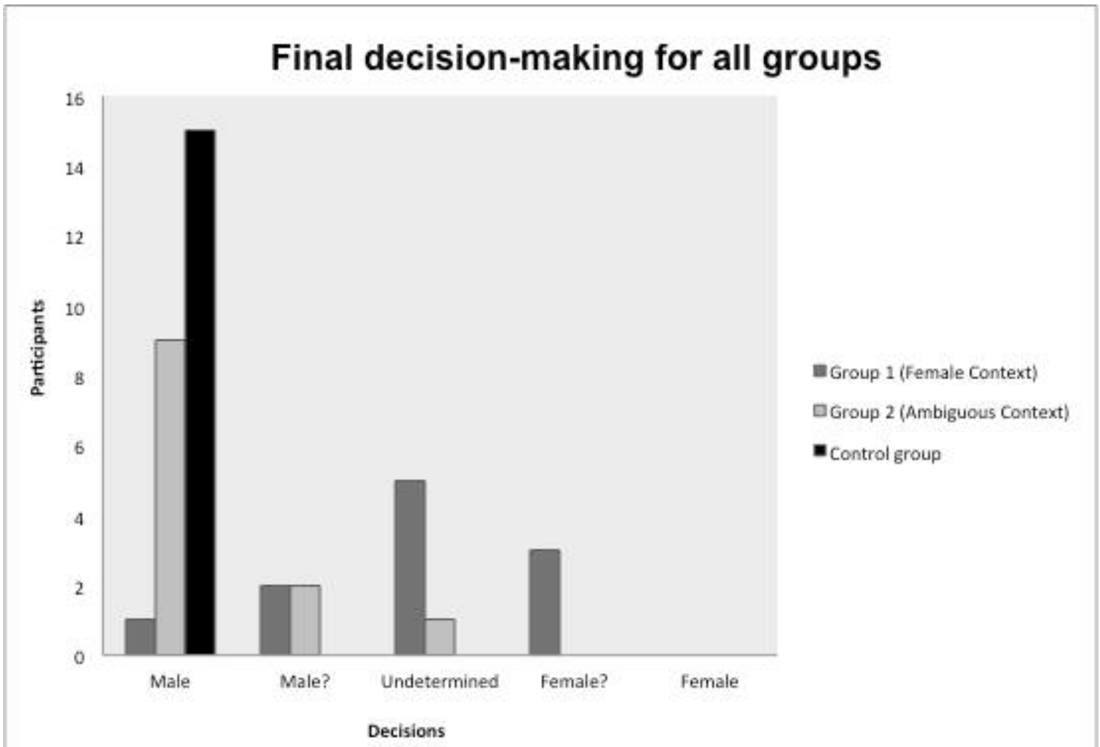

Figure 4.5 showing the distribution of the decisions on the skeletal remains for all three groups sex assessment

4.3.2. Chi- Square and Fisher Exact test comparing groups in sex assessment

Control vs. Group 1 and Group 2

The Chi -square and Fisher exact test was used to compare the Control group to both crime scene groups (Group 1 and Group 2) in order to see whether there was a significant difference between the groups as a function of the 'extraneous' contextual information. The result of the Chi-square test revealed a significant difference between Control group, Group 1 and Group 2 with a Chi-square, <0.005 and a p value of <0.01. The Fisher Exact test also showed <0.003 and a p value of <0.01. (see Table 4.2 for further details).

Control vs. Group 1

The Chi-square and Fisher exact test was used to statistically determine whether the distribution of categorical variables between the Control group (no context) and Group 1 (Female context) differed significantly from one and other. The result of the Chi-square and Fisher exact test revealed a significant difference with a Chi-square, 0.000 and a p value of <0.01and the Fisher Exact test showing <0.000 and a p value of <0.01. (see Table 4.2 for further details).

Control vs. Group 2

The Chi-square and Fisher exact test was used to statistically determine whether the distribution of categorical variables between the Control group (no context) and Group 2 (Ambiguous context) differed significantly from one and other. The Chi-square and Fisher exact test revealed no significant difference with a Chi-square, 0.121 and a p value of > 0.05 and the Fisher Exact test showing >0.075 with a p value of > 0.05 (see Tables 4.2 for further details).

Group 1 vs. Group 2

The Chi-square and Fisher exact test was used to statistically determine whether the distribution of categorical variables between the Group 1 (Female context) and Group 2 (Ambiguous context) differed significantly from one and other. The Chi-square and the Fisher exact test revealed a significant difference with a Chi-square, 0.007 and a p value <0.01 and the Fisher Exact test showing <0.003 and a p value of <0.01 (see Table 4.2 for further details).

Table 4.2 showing the results of the Chi-square and Fisher's exact test for sex assessments for all groups compared

	Value	df	Asymp sig.	Exact sig.
Control vs. Group 1 and Group 2 (N=38)				
Pearson's Chi-Square	12.887	3	.005	.002
Fisher's Exact Test	11.848			.003
Control vs. Group 1 (N=26)				
Pearson's Chi-Square	22.159	3	.000	.000
Fisher's Exact Test	21.976			.000
Control vs. Group 2 (N=27)				
Pearson's Chi-Square	4.219	2	.121	.075
Fisher's Exact Test	3.721			.075
Group 1 vs. Group 2 (N=23)				
Pearson's Chi-Square	12.046	3	.007	.002
Fisher's Exact Test	11.669			.003

4.3.3. Confidence level of participants

Crime scene Group 1 and Group 2

An independent samples t-test was conducted to compare the confidence level for participants in Group 1 (female grave context) and Group 2 (ambiguous context), The results indicated a statistically significant difference at the 95% significance level in the confidence level for Group 1 (M=58 SD=15.63) given female grave context and Group 2 (M=81 SD=12.21) given ambiguous grave context; t(18)=-3.719, p=0.002 (see Table 2)

Group 1 vs. Control

An independent samples t-test was conducted to compare the confidence level of participants in Group 1 (female grave context) and the Control group (no context) The results show a significant difference in the confidence level for Group 1 (M=58 SD=15.63) given female grave context and the Control group (M=79 SD=7) given no context; t(22)=-4.659, p=0.000 (see Table 2)

Group 2 vs. Control

An independent samples t-test was conducted to compare confidence level for participants in Group 2 (ambiguous context) and Control group (no context) The results show no significant difference in the confidence level for Group 2 (M=81

SD=12.21) given ambiguous grave context and the Control group (M=79 SD=7) given no grave context; t(24)=0.416, p=0.681 (see Table 2)

Control group

A paired sampled t-test was conducted to compare confidence level for participants in the control group before and after context. The results show a significant difference in confidence level before (M=79 SD=7) and after (M69 SD10) context; t(14)= 4.675, p=0.000 (see Table 2).

Table 4.3 An overall summary of the mean confidence value across all groups

Group	Mean (%)	SD	N
Group 1 (Female context)	58%	15.63	11
Group 2 (Ambiguous context)	81%	12.21	12
Control Group	79%	7.00	15
Control Group after Context	69%	10.00	15

4.3.4. Decision-making change of participants in Control group after context

After all 15 participants assessed the male skeletal cast blindly, 7 participants received the summary report of the female burial contexts and 8 participants received the summary report of the 'gender neutral' burial contexts. Only 2 participants in total (1 participant from each sub-group) changed their initial decision on the male skeletal casts after receiving the summary context. The initial assessments changed from male to undetermined and from male to male?

4.4. Discussion

Participants sex assessment

The findings of this study show that initial exposure to context at a crime scene can affect the subsequent assessment of the skeletal remains. The results indicated that there was a difference in the sex assessment made by the participants of the male skeletal cast that was highly dependent upon the context they were exposed to prior to the analysis. This was increasingly noticeable when participants were exposed to a 'strong' context. For example, in Group 1 (female context), only one participant (9%) assessed the skeletal cast to be 'male', two assessed it to be 'male?' (18%), three assessed it to be 'female?' (27%), with five of the participants providing an assessment of ‘undetermined’ in their final interpretation (45%). However, in Group 2 (ambiguous context), 9 participants (75%) assessed the skeleton to be ‘male’, two stated ‘male?’ (16%) and only one participant (8%) provided a conclusion of ‘undetermined’ in their assessment.

The cascading effect of the contextual information was also notable when comparing both groups to the control, with all participants in the control group assessing the skeletal cast to be male. This demonstrated that the female clothing associated with the male skeletal cast did affect the sex assessment of the skeletal remains, whilst the gender-neutral setting did not have as much of an affect upon the final sex assessment reached by the participants. This indicates that not all contexts in all situations will have an affect on the subsequent assessment of the skeletal remains and shows that ‘neutral’ clothing did not affect the interpretation of the skeletal remains. Previous studies addressing contextual influences and forensic anthropology have shown that when ambiguity is involved in the assessment of skeletal remains, a strong context (such as DNA) influenced the interpretation of participants with regard to sex assessments on the skeletal remains [26]. However, this study highlights the fact that what is considered as influential will depend not only on the nature of the task (Dror 2014a) but also on the level of ambiguity of the characteristic of the evidence being interpreted (the difficulty of the judgment), and the strength of the context in which the decision is made, as well as on the direction of the bias (Dror 2016).

Confidence level

The results also showed that there was a significant difference in the confidence level of the participants when comparing Group 1 and Group 2, with participants in Group 1 having a lower certainty in general in their assessment and final evaluations compared to Group 2. This indicates that although participants in Group 1 arguably did not make the ‘correct’ assessment of the skeletal remains of the male cast, their confidence level in the final assessment suggests that they were not as confident in their judgments when compared to participants in Group 2. Exposing participants to a strong ‘extraneous’ context (such as female clothing) might have created an early hypothesis and initial belief that the skeletal remains were in fact female. Studies have demonstrated that prior beliefs can be resistant to change (Burke 2005), and that once a hypothesis it formed it is difficult to adjust the tenacity of that belief even after receiving new information that contradicts or dis-confirms the basis of that belief (Anderson et al. 1980). Therefore, it is perhaps not surprising that the majority of participants in Group 1 were not as confident in the sex assessment of the skeletal remains as the context might have contradicted their initial belief of the skeletal remains to be of a female. Equally, for participants in Group 2, the gender neutral context might not have created as strong an initial belief compared to Group 1, and therefore resulted in participants having more confidence in making their final evaluations when making a sex estimation on a clearly male skeletal cast.

Similar results were found in the control group where the 15 participants also showed a difference in their confidence level before and after receiving the short report. Although participants did not change their initial judgment of the skeletal cast being a male, their confidence in that judgment was reduced after receiving the description of the clothing and grave artifacts associated with the skeletal remains. This indicates that context did indeed affect confidence in the decision-making, but not the decision-making 'outcome' of the skeletal remains.

Metric and non- metric assessments

This study included both metric and non-metric analysis on non-ambiguous skeletal casts, and it is important to highlight that the aim of this study was not to conduct a validation and classification study of non-metric and metric methods used in forensic anthropology. The focus of this study was to look further into the role of early exposure to context at a crime scene, and how that might unconsciously influence subsequent analysis at a later stage. This is important, as previous validation and classification studies within sex assessments in forensic anthropology have generally shown these methods to be reliable, with high classification accuracy, specifically for sex estimation of the pelvis (Klales et al. 2012; Ubelaker & Volk 2002) Furthermore, the assessment of the participants was based on the basis of the overall inferences made from all methods available, (both metric and non-metric), rather than on one technique, or the single traits scored for each method. However, the majority of decision-making 'uncertainties' for participants in Group 1 were based more within the non-metric methods used, as the results from the metric methods showed (according to the measurement), the skeleton to be clearly from a male.

Previous studies within forensic anthropology have shown that people tend to rely upon visual methods more frequently than metric ones specifically within sex assessments (Klales 2013). It is plausible to suggest that participants in this study tended to rely more upon the visual traits, giving room for interpretations more in accordance with their initial beliefs, as prior studies in forensic anthropology have shown that there is a tendency to change the scaling of single traits to fit the overall decision reached (Klales & Lesciotto 2016). Moreover, the morphological traits of skeletal casts are not as 'clear' in features as real skeletal remains, and therefore arguably an element of ambiguity on certain traits might have been inherent to the experiment, causing participants to unconsciously rely on the context further when making decisions on visual assessments.

Limitations

It is important to highlight that this study was based upon a mock crime scene, with a limited sample size (due to participant availability), with non-working experts within the field of forensic anthropology. Although being an expert has been shown to generally lead to higher performance, there are also cognitive vulnerabilities inherent in expertise due to the mechanisms of the brain for storing and processing information (Wood 1999; Dror 2011). A recent empirical study with experts in crime scene investigation showed that prior information did effect experienced crime scene investigators; they interpreted the crime scene differently dependent on the prior information that the examiners obtained (van den Eeden et al. 2016). Therefore, a valuable comparable study would be to see if similar effects could be found amongst working professional anthropologists.

Furthermore, studies have also identified issues for potential expectation bias when estimating sex of the skeletal remains (Effros et al. 2000; Walker 1995). However, it has also been highlighted within the domain of forensic anthropology that contextual information specific to a case is of high importance, especially in trauma assessments, where a lack of context could have a severe impact on the interpretation (Pinheiro et al. 2015; Blau 2016). Despite the discussion about the effect of context in forensic anthropology, further empirical studies are required in order to establish what factors and under what circumstances, and at what stages within the biological profile approach, context might influence the decision-making process and subsequently bias the interpretation of the skeletal remains.

4.5. Conclusion

This study has provided an important step towards understanding the potential effects of initial exposure to contextual effects at a crime scene upon judgment and decision-making within forensic anthropology. This study specifically showed that 'extraneous' grave clothing associated with skeletal excavations impacts upon initial beliefs, judgments and the subsequent assessment of the skeletal remains. This was increasingly noticeable when participants were exposed to a strong female context. Furthermore the results also showed that there was a significant difference in the confidence level of the participants, depending on the context. Similar results were found in the control group where a difference in confidence level of participants was identified before and after receiving context.

Chapter 5. The order of examination on skeletal remains and cognitive bias

5.1. Introduction

The findings presented in experimental chapters 3 and 4 indicate that contextual information and early exposure to a strong extraneous context can affect the analysis and interpretation of the skeletal remains when establishing a biological profile. However, the general conditions that influence the extent of cognitive bias may also be shaped by multiple factors aside from contextual and extraneous stimuli influences (Dror 2013). Indeed, it has been demonstrated/suggested that the working environment itself and the procedural practice within a forensic discipline may also impact the degree to which cognitive biases affect evidence interpretations (Dror 2013). Some have suggested that forensic work should be conducted linearly, meaning that evidence needs to be examined and analysed in isolation from a target comparison (Dror et al. 2015), in order to avoid any sources of influences that could compromise the interpretation of the evidence. Failing to conduct a linear approach may consequently lead to the target influencing the analysis of the evidence (Dror 2009).

Similarly, mislabelled, interchanged, and contaminated samples could create numerous possible problems, within the evidence itself as well as the interpretation of the evidence (Koehler 2016), including the possibility of cognitive biases entering the investigation at an early stage. For example, within recent years, the integrity of laboratories has been called into question, with some controversial cases highlighting the lack of standards and absences of quality control measurements resulting in questionable evidence being used in court (PCAST 2016). The lack of high quality standards have lead to closing of laboratories, such as the Detroit police crime lab where audits found erroneous findings in ten percent of two-hundred random cases, where forensic evidence was used to prosecute, due to the mishandling of forensic evidence (Bolton-King 2016). Furthermore, the 2009 NAS report as well as 2016 PCAST report also highlighted the issues of problems with inconsistent practice in crime laboratories, urging forensic science disciplines to create further accreditations and standards within the different professions.

As a response many forensic domains are now establishing guidelines, standards, and accreditation processes in order to help guide the conduct of best practice, as well as minimizing any biasing effects. As mentioned previously, (Chapter 2 section 2.3) in forensic anthropology, professional working groups such as SWGANTH, have developed consensus best-practice guidelines and established minimum standards for the forensic anthropology discipline. Although, SWGANTH highlights that the anthropological tests should be preformed without external influences (SWGANTH Laboratory Management and Quality Assurance, 2011, p.4), very little is known in forensic anthropology with regards to cognitive biasing effects beyond external contextual influences, that could in fact alter the interpretation of the skeletal remains, based on how one conducts a biological profile.

As highlighted in the literature review, (Chapter 2 section 2.3.9) it is necessary for the work of Weiss (1972) and Walker (1995) to be expanded upon specifically regarding the possibility of the standard practise of sexing methodologies offering opportunities for contextual bias due to the skeletal remains acting as a biasing context itself. This is important not only due to further improving the creditability of standard procedures within the practice, but also to avoid any potential biasing snowball effect where one piece of information, potentially biasing another element within the identification process (Dror 2012). As mentioned previously, sex estimation is one of the first steps in the process of a positive identification of an unidentified individual (Gyomarch and Brizek 2011). This is due to the fact that many methods applied for age at death, ancestry, and stature being sex specific (Klales 2013), especially within visual methods. This could potentially mean that that if there are cognitive interpretation issues arising during the stage of sex assessment, arguably the interpretation of age at death could be exposed to biased evaluations as well, as a result of a domino effect. This chapter will therefore exclusively focus on visual sex assessments within forensic anthropology and address whether the order of sex examination of skeletal remains could influence, a) the interpretation of the subsequent skeletal element (i.e. if examining a clearly male pelvis will consequently skew the interpretation of the skull morphology and vice versa) and b) if the order of examination will determine the final conclusion of the sex of the remains.

5.2. Methodology

5.2.1. Experimental design

Experimental studies were designed to test whether the order of examination of the os coxa and the skull could act as a biasing context**.** More specifically the design of the experiment was set up to see if the sex assessment of the skull/os coxa could serve as a biasing ‘context’ on the subsequent analysis depending on the order in which the skeletal remains would be examined. This was done by dividing participants in two groups where one group first assessed the sex of an innominate, then assessed the sex of a skull, (using visual methods standard in the field), and the second group assessed the sex of the skull, following assessing the sex of an innominate (see Appendix D).

In this study, the skull presented very strong female morphology and the os coxa presented very strong male morphology, having a skull and a os coxa taken from two different individuals. However, in the briefing of the participants, it was implied that the paired elements were from the same individual. This was done in order to assess the degree to which the order of examination skewed the scoring of the morphological sex traits on the following skeletal element, resulting in a difference in the final conclusions regarding the sex of the remains by the participants. Participants were also asked to give a confidence level in their conclusions in order to gain an insight into the confidence of each participant in their decision-making. In addition, participants also answered follow up questions (see Apendix D), which gathered demographic background information of the participants.

5.2.2. Material

One skull and one innominate were selected from the archaeological collection provided by the UCL Institute of Archaeology. The skull and the innominate were chosen from two separate individuals in order to be able to provide a skull and an os coxa showing strong contradicting morphological features. These exhibits were chosen in this way in order to assess whether the order of examination of the skeletal elements would lead to a bias on the following assessment and the final interpretation of the participants. This resulted in a selection of a skull from an individual (previously assessed by osteologists) with a classical female morphology, meaning, (very generally speaking) the skull to be very ‘gracile’ with a smooth and more vertical frontal bone, small mastoid processes, a small maxilla and a sharp supra orbital margin. Furthermore the innominate bone from one individual (right os coxa) with typical male characteristics was also chosen, including traits such as a v-shaped sub-pubic angle, narrow greater sciatic notch with a high and vertical ilium. The left side of the innominate (os coxa) was badly damaged; therefore, participants were only shown and asked to assess the right os coxa. Furthermore, the colouring of the skeletal elements was of similar nature, which was important in order to imply to the participants that the paired skeletal remains may have originated from the same individual. See Figure 5.1. and Figure 5.2

Figure 5.1 showing the right os coxa used in the experimental study

Figure 5.2 showing the skull used in the experimental study

5.2.3. Participants

The participants chosen were MSc and PhD students with prior experience in sexing skeletal innominate and skull using morphological techniques derived from osteological sexing methods. Thus, participants were recruited from an emailing list of former and current students in Forensic Anthropology and Bioarcheology (see Appendix D). In order to not compromise the true nature of the study, participants were told that the focus of the study was to document the confidence level in using sexing techniques on skeletal remains based on visual assessments. Furthermore, demographic information was collected for each participant in order to gain further insight into the background knowledge of each participant, and their practical experience in using visual sexing techniques. Participants were asked, in this experiment, to score on the traits outlined mainly by Phenice (for the os coxa), and Walker/ Bukistra and Uberlaker for the skull. Additionally, participants were also given access to books and journals demonstrating the methods, in addition to any personal notes on sex assessments that they might have obtained prior to the assessment. All participants provided informed consent, and all data were anonymised following standard data protection protocols (data protection act 1998) (see Appendix D).

5.2.4. Procedure

Participants were semi-randomly divided into one of two groups, Group 1 and group 2. Participants in Group 1 were asked to first assess the sex of the skull, followed by the os coxa, with participants in Group 2 being asked to assess the sex of the os coxa, followed by the sex assessment of the skull. All participants assessed the same skeletal elements with the only difference being the order in which the skeletal remains appeared. Participants conducted the assessment alone with only a senior forensic anthropology lecturer present to provide participants with the materials. Each analysis of the skeletal remains was conducted in 'isolation' and participants were only given the second skeletal element after completing the first assessment. This was to ensure that the participant carried out the initial assessment in isolation from any further influences.

Each participant was given as much time as needed, however the average time for a participant to complete the assessment was approximately 20 minutes. Participants recorded their answers on an answering sheet with each participant given a confidence level for each of their assessments. This was also followed up with a questionnaire form to collect further information pertaining to the background and experience of each participant (see Appendix D for further information). The questioner form included questions such as, what method participants were most familiarized with during their education (visual vs. metric), as well as which skeletal element(s) participant preferred to use when conducting a sex estimation. This was done in order to further understand how familiarized participants were with visual methods, as well as which method and skeletal element in sex assessment participants preferred and therefore possibly put more 'weight' on when assessing sex of an individual.

5.2.5. Analysis

The data were recorded and analysed using descriptive statistics and SPSS for significant tests. Due to the data being categorical, with a small sample size in sex assessment a Chi-square and Fisher exact test was, used to examine whether there was a significant difference between the groups sex estimation on the skeletal remains as a function of the order of examination. Furthermore, a series of independent t-tests were conducted to assess whether the means of the two groups confidence level statistically differed.

5.3. Results

Table 1 presents the demographic and background summaries of the fifteen participants.

Table 5.1 showing background summaries and demographic for the participants

Variable	*n*
Sex	
Male	2
Female	13
Highest level of education	
MSc	11
PhD	4
Educational background	
Archeology	4
Bioarcheology/Forensic anthropology	7
Osteology	3
Physical anthropology	1
Years of practice in forensic anthropological/osteological methods	
1-2	5
2-4	4
5-7	6
Focus on method(s) training during education	
Metric	0
Non-metric	4
Both	11
Preferable method(s)	
Metric	4
Non-metric	4
Both	7
Preferable skeletal element(s) to use when assessing sex	
Pelvis	10
Skull	1
Pelvis and Skull	4

5.3.1. Participants decision-making on the Skull and the Os coxa

Figure 5.3 illustrates the decisions made by each participant in Group 1 and Group 2. The results demonstrates all eight participants in Group 1 correctly assessing the skull to be female, and all seven participants in Group 2 also accurately assessing the skull to be female. In Group 1, six participants correctly assessed the os coxa to be of a male with two participants assessing it to be undetermined. In Group 2, all seven participants assessed the os coxa to be of male.

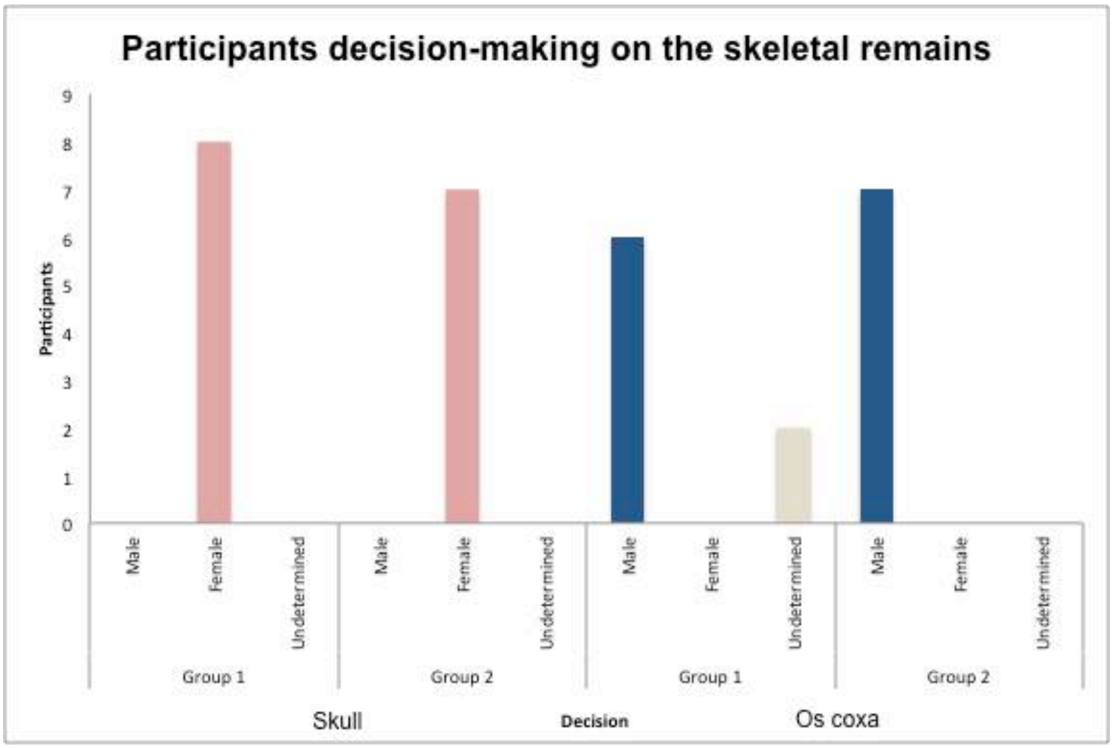

Figure 5.3 showing the interpretations made by each participant in Group 1 and Group 2 on the skull and the os coxa

5.3.2. Chi-square test comparing final sex assessment for Group 1 and Group 2

The Chi-square and Fisher's exact test revealed no significant difference in participants final sex estimation and order of examination between Group 1 (skull to os coxa) and Group 2 (os coxa to skull) with a Chi-square > 0.622 and a Fisher's exact test > 0.622, showing a p-value > 0.05. Figure 5.4 shows the distribution of participants' final sex estimation on the skeletal remains for both groups.

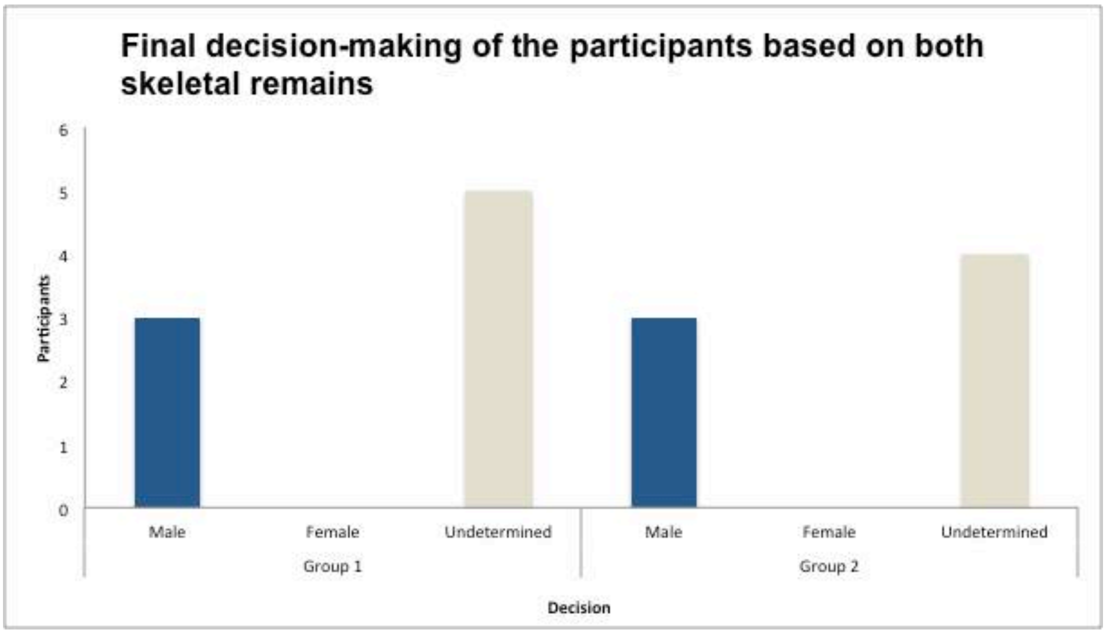

Figure 5.4 showing the distribution of participants' final sex estimation on the skeletal remains for both groups

5.3.3. Independent samples t-test Os coxa

An independent sample t-test was conducted to compare the means of the two groups based on the confidence level stated by each participant, in order to see if there was a significant difference in participant confidence in their sex assessment of the os coxa based on the order of examination.

The results showed that there was no significant difference between confidence levels between Group 1 (Mean=79,SD=10) and Group 2 (Mean=87,SD=9), depending on the order of examination, t;(13)=1.574,p=.140.

5.3.4. Independent samples t-test Skull

An independent sample t-test was conducted to compare the means of the two groups based on the confidence level stated by each participant, in order to see if there was a significant difference in participant confidence in their sex assessment of the skull based on the order of examination.

The results showed that there was no significant difference between confidence levels between Group 1 (Mean=79,SD=14) and Group 2 (Mean=70,SD=17), depending on the order of examination, t;(13)=1.168,p=.264.

5.3.5. Independent samples t-test final assessment for Group 1 and Group 2

An independent sample t-test was conducted to compare the means of the two groups based on the confidence level stated by each participant, in order to see if there was a significant difference in participant confidence in their final sex assessment based on the order of examination.

The results showed that there was no significant difference between confidence levels between Group 1 (Mean=79,SD=13) and Group 2 (Mean=70,SD=12), depending on the order of examination, t;(13)=1.338,p=.204.

Table 5.2 An overall summary of the mean confidence value across the skeletal elements for each group

Skeletal element	Mean (%)	SD	N
Os coxa			
Group 1	79	10	8
Group 2	87	9	7
Skull			
Group 1	79	14	8
Group 2	70	17	7
Final assessment			
Group 1	79	13	8
Group 2	70	12	7

5.4. Discussion

The findings of this study show that the order of examination in sex assessments of the skull and the os coxa did not affect the overall sex assessment on the individual by the participants. The results indicated that the majority of participants correctly assessed the skull and the os coxa in isolation, as well as combined, with no indication of an influence, regardless of the order of appearance, in which the skeletal remains were assessed. Furthermore, the results did not indicate any difference in confidence level in sex assessments depending on order of examination between the groups, when comparing both skeletal elements in isolation as well as combined.

However, although the majority of participants in both groups 'correctly' assessed the skeletal remains in isolation as well as combined, forty percent of all participants across both Group 1 and Group 2 gave more weight to the precise sex of the os coxa than the precise sex of the Skull, when making a definite overall sex estimation. This resulted in six out of fifteen participants estimating the individual to be of a male. As illustrated in figure 5, the majority of participants (sixty percent) did estimate the sex of the individual to be ‘undetermined’ based on the mixed traits and contradicting morphology presented by the skull and the os coxa, stating that further analysis is needed in order to make a definite conclusion (see Appendix D for follow up questionnaire). It is possible to infer that the participants, who concluded that the overall sex of the individual was ‘undetermined,’ reached a ‘correct’ decision when dealing with limited skeletal remains with mixed morphology from one individual. However, sexing based on the features of the innominate is considered superior in forensic anthropology, (compared to the skull and post-cranial elements), with standard procedures in the field determining the innominate to be more reliable when assessing sex of an individual (Spradley and Jantz 2011, Kimmele et al. 2008, Meindl et al 1985, Rogers, 2005, Walrath et al. 2004 and William and Rogers, 2006). Furthermore, when asked in the questionnaire form which sexing method participants preferred using when assessing sex, majority of the participants (sixty-seven percent) said the pelvis bone. Therefore, the results of the participants who reached a definite conclusion of the individual being of a ‘male’ does not necessarily constitute an incorrect answer, as the standards in the field allow for such decision-making based on the reliability of the method applied. This does however raise concerns with regards to the possibility of valuable information from the skull being ignored and not fully taken into account, if the innominate is present. It questions the role of the skull in standard sex estimations and indicates that more research within this area is needed in order to reconsider its value.

Although the debate on the use of sex estimation of the skull and the reliability of the morphological traits has been questioned (Weiss 1972 and Walker 1995), the methods used in sex assessments of the skull are still widely held by the general forensic anthropology community to be reliable (Rogers 2005). The challenge for the forensic anthropology community is to identify how this could potentially affect the interpretation of a) forensic cases where only the skull might be present, and b) potential misidentifications of commingled burials if a mismatched innominate is associated with a skull. This is because commingling of human remains is a common component of certain types of incidents such as, mass disasters, terror attacks, or mass grave excavations, which often involves a multidisciplinary approach to a complex forensic investigations (Uberlaker 2014). The nature of the incident may result in extensive fragmentations as well as comingling of human remains, presenting special problems for the identification process (Steadman et al 2014).

Limitation

There were some limitations in this study, due to the small sample size of the participants, and also the design of the experiment. The time limit of conducting this experiment as well as getting participants to come in-person and conduct the study restricted the sample size of the participants. Furthermore, the skeletal elements used in this study were too distinctly male or female. Cognitive biases in forensic research have been shown to be more prevalent when dealing with decision-making under uncertainty (i.e. distorted finger-marks, mixed DNA samples, ambiguous skeletons etc.). A valuable comparison in this study would have been to assess whether similar observations would be made if the order of examination had included one clear/distinct skeletal element followed by an ambiguous one. It would also be interesting to see under what conditions the innominate may act as an influence on the assessment of the skull. This could be done by providing one group of participants with a female innominate followed by an ambiguous skull, and the second group with a male innominate followed by the same ambiguous skull, in order to see if the scoring of the morphological sex traits of the skull would differ depending on the innominate associated with the skull.

In addition, this study did not include other variables such as ancestry and age at death estimation, which could act as influencing factors when conducting a biological profile. As mentioned in the literature review, forensic anthropological methods need to be evaluated within the population being studied. As discussed in Chapter 2 section 2.3, studies have shown that there is a tendency for example to misclassify Hispanic males as females, due to the appearance of their gracile morphology (Spradly et al. 2008). Furthermore, studies have also shown misidentification of archaeological remains, where elderly female skulls have been mistaken for male skulls, due to their robust appearance, related to the aging of the skeleton (Weiss 1972; Walker 1995). This also suggests that future studies including all factors that potentially could create a biasing impact upon the sex assessment of the skeletal remains should therefore be evaluated. Furthermore, this should not only be considered within visual assessments but also within metric ones.

5.5. Conclusion

The results of this study show that the order of examination in sex assessments of the skull and the os coxa did not affect the final decision-making of the participants. The participants correctly assessed the skull and the os coxa in isolation, as well as when assessing them together in a combined scenario. Furthermore, the results also indicated that participants give more weight to the precise sex of the os coxa than the precise sex of the skull, when making a definite overall sex estimation. This does however highlight that further research is needed in order to re-assess the 'value' of the skull as a sex indicator for forensic cases. This is especially important when dealing with forensic anthropological cases where the skull is the only skeletal element present, or the innominate being badly damaged, as well as when dealing with commingled remains. The study conducted in this experimental chapter has only begun to investigate the potential for cognitive interpretation issues involved in the methodological approach and procedure in sex assessments with additional research needed in order to understand its implication further in biological profiling.

Chapter 6. Discussion

This thesis had three primary research aims. Experiment 1 and 2 (Chapter 3) explored empirically whether contextual information can affect previous judgments when assessing skeletal remains of ambiguous nature. Experimental study 3 (Chapter 4) represented the first empirical test of the possible effects of early exposure to extraneous context at a crime scene on the subsequent assessment of the skeletal remains. Finally, experimental study 4 (Chapter 5) studied whether the order of examination of skeletal remains could influence the interpretation of the subsequent skeletal element, as well as the final conclusion. The implications of each of these studies is discussed specifically within the specific domain of forensic anthropology, as well as within the broader forensic science. How these findings contribute to the existing body of knowledge and research within forensic science, and future avenues for further studies are then considered.

6.1. Visual assessments and contextual effects

In a manner akin to previous studies regarding the effect of context on the interpretation of forensic evidence (Nakhaeizadeh, Dror et al 2014), the findings from experiments 1 and 2 highlight and extend the findings that visual assessments in forensic anthropology can change as a result of contextual information. Experiment 1 and 2 not only show that visual assessments are mediated by context, but can even change and override the previous assessments made by participants on the same skeletal remains. The literature has repeatedly shown that judgments can be affected by information that suggests the possibility of a particular outcome (e.g. Bieber 2012). Moreover, psychological research has well established that top-down processes can have a fundamental affect on visual perception and bottom-up information (e.g. Balcetis and Dunning 2006. Thus, some researchers suggest general caution with regards to any 'information' that is not essential for the given analysis (e.g. Dror et al. 2015). The results of the two experimental studies (experiments 1 and 2) indicate that the 'suggestive' contextual information provided to the participants (on previously assessed skeletal remains) caused them to re-evaluate and alter how they perceived the evidence.

The findings emphasise a key argument in favour of 'shielding' examiners from case information that is not pertinent for the given analysis, namely because the context changed and possibly also 'undermined' the independency of the examiners initial analysis of the skeletal remains. Further, the two experimental studies show a lack of consistency in the interpretation of the skeletal remains when context is given, which has previously not been considered within the domain of forensic anthropology. Although the two studies did not include any professional anthropologists as participants, the inconsistency in the decision-making of the participants when given context still offers some insight into factors that impact human performance.

The findings also highlight the need for further research within reliability issues in terms of addressing the consistency of expert performance and how this relates to potential 'bias' within and between experts when context is given in the assessment of skeletal remains (see Dror 2016). Together, studies 1 and 2 thus extend the findings of other studies in other domains (Dror et al. 2005; Dror et al. 2006; Dror & Hampikian 2011) with respect to the impact of context in forensic decision-making, especially when those decisions are complex and difficult. This further exposes a vulnerability to contextual influences in subjective methods used in forensic anthropology.

The results of study 1 and 2 also raise important questions as to the exact nature of the process whereby contextual information overrides previous judgments. Hence, these experiments not only show that context can mediate judgments, but that it can potentially cause people to override and contradict their previous decisions. It highlights the concerns raised by the National Academy report (2009) and The PCAST report (2016) with regards to the issues of potential contextual biases that might result from methods that are reliant upon human judgments. One might argue that participants in experimental study 1 and 2 were not 'biased' due to the mere fact that no absolute 'ground truth' about the ambiguous skeletal remains used could be established, and therefore no 'judgment accuracy' could be measured. Nevertheless, the studies did show that contextual information altered the way in which the skeletal remains were interpreted. For example, the results from experimental study 2, showed an overall change to the previous assessment made by participants (63%) across the decisions made, with the majority of the participants making a decision that confirmed the contextual information given. This demonstrates that exposure to contextual information could possibly produce cognitive biases.

In forensic anthropology the environmental and contextual influences will vary depending on the legislative practice and the procedures being followed. The practice of providing forensic anthropologists with access to case information will also therefore vary depending on the aforementioned reasons. However, it is not uncommon for forensic scientists in general to have access to case reports, (e.g. pathology reports, police statements, laboratory analysis etc) that could indeed be extremely influential. Although experimental studies 1 and 2 did not include any ‘forensic case reports’, much can still be drawn from the issues that may arise when relying on the context when assessing skeletal remains.

Firstly the findings show that even relatively ‘weak’ context can influence the decision-making outcome, as the context used was not considered to be ‘overpowering’ or ‘emotionally loaded’ compared to other studies (e.g. Nakhaeizadeh et al 2014a and Dror et al 2005). Second, the majority of participants changed their own previous evaluation of the skeletal remains, (a re-evaluation so it fitted with the contextual information given), thereby showing clearly that context does not need to be ‘extreme’ in order for it to have a strong affect on visual methods. Third, this raises concerns with regards to possible sources of contextual biases that might enter in forensic anthropological casework. Many of the cases in forensic anthropology involve the identity of an unknown deceased person, where the police may have an idea of the identity of the remains by searching missing persons databases (Marquez 2015). Consideration should be taken (based on the result from experimental study 1 and 2 of this thesis) of how that information might affect the decision-making process in the establishment of the biological profile, specifically in cases where the context might be suggestive if not overpowering. Knowing what information to obtain, and what questions to ask, has been argued to save considerable time and effort in forensic anthropology (Konigsberg et al. 2009). One of the most poignant dilemma for forensic anthropologists arises when they have to decide whether (and what) information to request, in order to make a positive identification. The results from study 1 and 2 presented in this thesis however, suggest that information does alter the interpretation of skeletal remains, and therefore consideration must be made as to how to identify and counter possible sources of contextual influences. This should particularly be considered when dealing with incomplete or distorted skeletal remains, (which is not unusual in forensic cases) where the risk of unconsciously relying on police statements and other information might be more likely, as additional information in complex cases is more sought to contribute to the decision-making and inference process.

Finally, the results also highlight the possibility of what is described by Dror (2012) as the 'bias snowball effect' (see also what Kassin (2012) refers to as "corroboration inflation"). The results from studies 1 and 2 demonstrated a context effect and a change to a previous interpretation of skeletal remains across sex, ancestry and age at death. As explained in chapter 2 (section 2.3.3), most visual assessments are sex dependent, meaning that if there are cognitive interpretation issues arising during the stage of sex assessment, arguably the interpretation of age at death could be subject to 'biased' evaluations as well. This could potentially cause a domino effect that not only causes interpretation issues during the different analysis of the skeletal remains, but may also affect the way in which evidence is presented in court. Preferably, multiple morphological features are to be used in assessing skeletonised remains, although this may not always be possible in forensic contexts. Skeletal remains could be damaged due to poor preservation, and in some cases even burned and fragmented, resulting in highly ambiguous scenarios. Forensic case scenarios become even more challenging when taking into account cognitive processes involved in complex decision-making. The results of study 1 and 2 show that one must be aware of the cognitive limitations involved in each visual assessment in order to avoid a potential biasing snowball effect where one piece of information, potentially biases another element within the identification process.

6.2. Initial exposure to context at a crime scene

There is a growing consensus in the forensic science community with respect to the existence of cognitive biases and the need for context management (e.g. Mattijssen et al. 2016; Dror et al. 2015). As mentioned in chapter 4, the most common solution proposed is the creation of a context controlled environment; in which the analyst is separated and blinded from potentially biasing information. This is important because the opinions of forensic scientists are considered impartial and unbiased, and not influenced by elements of the case that have no relevance to the scientific process

(Dror 2014a). The findings from Experimental 3 showed that ‘extraneous’ grave clothing associated with skeletal excavations had an impact upon initial beliefs, interpretations, the confidence level of judgments, and the subsequent assessments of skeletal remains.

Unexpectedly experiment 3 showed that early exposure of a strong indication of a certain outcome did in fact ‘bias’ the interpretation of non-ambiguous male skeletal casts. In contrast to previous research where cognitive and contextual effects have been shown to be more prone and make an impact when making decisions on ambiguous skeletal remains, or ‘difficult’ and challenging finger marks or DNA (e.g. Nakhaeizadeh et al. 2014a; Nakhaeizadeh, et al. 2014b; Dror et al. 2011; Dror et al. 2006), this study shows that contextual effects can also have an impact on non-ambiguous evidence. The results also highlight that not all contexts in a given situation appear to affect the final assessment made. For example, the findings only showed a bias cascading effect on subsequent assessment when exposed to a very strong context (female clothing on male skeletons). This shows (compared to experiment 1 and 2) that when dealing with non-ambiguous skeletal remains, only a very ‘strong’ context altered the interpretation made by the participants when assessing the male skeletal cast. Moreover, it highlights that the level of influence of contextual information will depend on:

- The nature of the task (Dror 2014a) and the ambiguity level of the given characteristic of evidence being interpreted (the difficulty of the judgment).
- The strength of the context in which the decision is made, as well as on the direction of the ‘bias’ (Dror 2016).

There have been some solutions proposed that seek to shield forensic experts from being ‘biased’ at an early stage of an investigation. Some have suggested a separation of different roles whereby the expert collecting the evidence not necessarily being the same one analysing it in the laboratory (e.g. Kassin et al. 2013; Saks et al. 2003; Krane et al. 2008; Dror 2014b; Dror et al. 2015). In forensic anthropology it is important that the forensic anthropologists are present on site in order to help preserve, excavate, and document the skeleton in *situ,* and mitigate the potential for the loss of important information pertinent to the anthropological assessment of the remains. Whilst it is recognised that it is important to utilise a combination of different types of evidence in the creation of a biological profile, this carries the risk of the anthropologist being exposed to 'extraneous information' at a very early stage of a forensic investigation. This needs to be considered and measured when developing approaches for scene management, evidence collection, and assessment. It is also important to consider that grave artifacts and items of clothing associated with skeletal burials are evidence in their own right. The results from experimental study 3 showed that clothing was influential. The results therefore indicates that it may be beneficial for these items to be considered separately from the assessment of the bone features to reduce the potential for cascading bias, which may arise as a result of irrelevant information cascading from one stage to another (Dror et al 2017).

Further, most every forensic case involves a variety of different specialised personnel, with both scientists and law enforcement working closely together, and thus a mixture of skillsets. Addressing and removing 'irrelevant context' has therefore raised concerns that such an approach may create silos of different personnel that hamper an integrated approach within the practice of forensic science, and in the delivery of robust forensic reconstructions. In many cases it is acknowledged that contextual information will have a role in assisting in forensic reconstructions. Whilst not all contexts will have a biasing influence, experimental study 3 from this thesis does illustrate that it is possible for context to affect early hypothesis and decision-making, even with non-ambiguous evidence. It is therefore important to be aware of such instances, and to take steps to ensure that inferences are shielded as much as possible form potential cascaded bias from the exposure of context.

Finding an appropriate balance between the risk and benefits of enacting solutions that seek to deal with the issues of extraneous context is not an easy undertaking. It could be argued that ideally, the forensic anthropologist collecting the skeletal remains might need to be different from the analyst conducting the biological profile, in order to allow the analyst to carry out their assessment without context associated with the death scene or the body itself. This would mean a change in working practices that may not always be feasible or straightforward, but this approach has been successful in some laboratories within other disciplines (Kassin et al. 2013; Krane et al. 2008; Dror 2014a).

6.3. Procedural practice and contextual influences

Compared to experimental studies 1, 2, and 3, experimental study 4 explored a different type of contextual influence. The generic conditions that influence the extent of cognitive bias may also be shaped by multiple factors aside from context. (Dror 2014a). Experimental study 4 was therefore designed to test for other 'types' of 'influences' that may arise due to certain procedural practice. Although the findings of experimental study 4 did not show such influences, the results did however show that some of the participants tended to rely more on the precise sex of the os coxa rather than the precise sex of the skull, when making a definitive overall sex estimation (40%). This raises some potential limitations concerning the value of the skull as a sex indicator for forensic cases, especially in events where the skull is the only skeletal element present, or the innominate being badly damaged, as well as when dealing with commingled remains. Although the data from experimental study 4 cannot address any of the questions and concerns raised, it can still provide some insight into potential issues that may arise.

As mentioned previously, one of the main challenges in laboratories when analysing commingled remains is to consider whether or not the remains originate from more than one individual are potentially intermixed (Marx et al. 2014). In some cases DNA profiles are provided as a powerful means of segregating remains, however this is dependent upon the resources available. Therefore, visual 'matching' of different element types is often used as one of the many means of identification (e.g. Byers 2010). Although guidelines recommend avoiding sole use of visual pair matching (due to the subjective nature of the procedure), visual matching may still occur, specifically if it involves a limited number of individuals. Consequently, the value and use of the skull in association with the innominate might possibly need to be further evaluated; in case a mismatched innominate is associated with a skull. This could potentially happen when dealing with complex scenarios such as mass disaster, and mass grave excavations, resulting in wrongful identification of the deceased.

6.4. Contextual effects in forensic anthropology

The context sensitive nature of each forensic case means that human interpretations are highly important, valuable and necessary. Humans are still needed to interpret results of sensitive and accurate analytical techniques, and to classify and identify evidence within the forensic science process. This creates complexities and controversy regarding how to best deal with human factors that could cause interpretation issues. The data from the experimental chapters presented in this thesis empirically show that context influences the undertaking of visual methods in which human judgment plays a central role. The studies established that the power of contextual influences in the assessment of the skeletal remains will differ, and in some cases result in contextual biases. The results from this thesis suggests that participants in this study (and potentially future forensic anthropologists) may not always be aware of where context may have a significant effect, and how it is being used and/or how it influences the interpretation of the skeletal remains, creating a lack of transparency in the decision-making process.

Therefore, the forensic anthropological community needs to start to differentiate between context that supports analysis of other evidence, and context that is evidence in its own right. For example, studies 1, 2 and 3, showed that participants tended to take the context as part of the analysis of the skeletal remains and not as an independent form of evidence (see Figure 6.1). This resulted in focusing on visual traits supportive of the context (experimental studies 1 and 2), changes in previous decisions (experiment 1 and 2), and contextual bias evaluations of non-ambiguous skeletal remains (experiment 3). This could arguably be due to a lack of a distinct inference and decision-making framework and approach within the discipline, showing just how evidence can be inter-related to the hypothesis in question, and the role of contextual influences within that process (Smit et al. 2016).

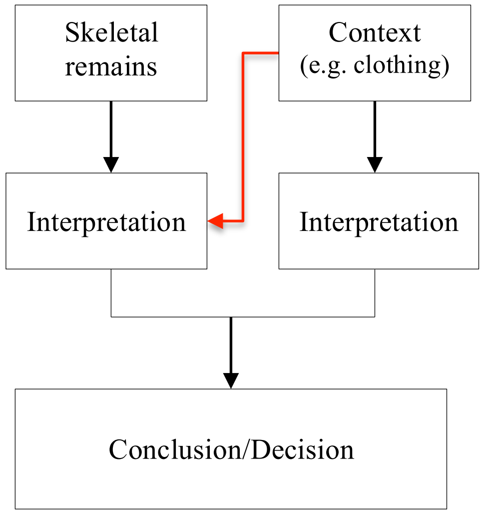

Figure 6.1 Showing the process of deciding upon biological profile based upon the skeletal remains and the context. Ideally the skeletal remains and context are interpreted separately (black line arrows). However the result from experimental study 3 shows that the context is being used in the interpretation of the skeletal remains (red arrow).

Even though the methods used in visual assessments in forensic anthropology are considered 'foundationally valid' and in principle reliable, there are still many reasons and factors that could affect examiners interpretations of the result. One of these reasons could be the deficiency in acknowledging the role and extent of cognitive and contextual influences within the discipline. Thus, it is important for the forensic anthropological community not to underestimate and minimise the importance of these issues as it has not only shown to affect expert interpretations across numerous forensic disciplines (e.g. Dror et al. 2006; Dror & Hampikian 2011; Osborne et al. 2014; Stoel et al. 2014; Nakhaeizadeh et al. 2014; Page et al. 2012; Archer & Wallman 2016), but to also affect the human role at the different stages of the forensic science process. Recognizing the role that cognition plays in the collection, analysis, interpretation, and presentation of evidence will enable the forensic anthropological community to address the concerns with regards to the issue of interpretations raised by reports such as the National Academy of Sciences report (2009) and PCAST report (2016), in the United States, and the Law Commission (2011) in the United Kingdom.

Another issue is that the parameters regarding what is considered best practice varies amongst forensic anthropologists, (in a similar manner to other forensic disciplines), where the handling of evidence at crime scenes and within laboratories diverges noticeably between countries and jurisdictions. The result from the experimental studies in this thesis have further highlighted that it will be important to offer approaches that can be sufficiently generalisable across different investigations and sample examination, but that are also sufficiently context sensitive to each case and each sample within it. It is widely recognised that every crime scene is different and it is therefore imperative to incorporate context sensitivity when looking to establish protocols for each discipline. This could be done by further generating empirical data that are indeed adequately generalisable but offer sensitivity to individual case context.

There has been a lot of recognition within the literature with regards to the need for empirical studies (Tangen et al. 2011; Petraco et al. 2012; Mnookin et al. 2011; Morgan et al. 2009). More recently, it has been recognised that sound research rather than training and experience, must become the central method by which assertions are justified (Mnookin et al. 2011), along with problem solving approaches that takes into account context sensitivity of each case (Margot 2011; Crispino et al. 2011; Ribaux et al. 2015). Notably, the PCAST (2016) report highlighted that forensic practitioners cannot rely on experience and extensive casework as a substitute for empirical studies of scientific validity. This is arguably due to the fact there is a lack of ground truth when approaching casework and therefore measuring validity and reproducibility will be ever more difficult.

The experimental studies in this thesis did not aim to measure the validity of the methods used in forensic anthropology, nevertheless these experiments have generated new data that will allow for a better understanding of decision outcomes and influencing factors in the forensic anthropological methods addressed. This is a significant contribution to the body of knowledge in terms of highlighting potential limitations of methodological approaches in a forensic reconstruction context, and also establishing which factors may influence the accuracy, reliability, and reproducibility of these approaches. This has not ever been empirically tested within the forensic anthropological domain, and the provision of data that establishes an evidence base of the extent to which these factors may be influencing decision making, is the best means of identifying the best steps forward.

6.5. The culture of error and research in forensic science and forensic anthropology

Within forensic science, the discussion of error has gained momentum in the wake of the NAS (2009) report, PCAST (2016) report, and forensic errors involved in wrongful convictions (Innocence project 2017). The term error however has a different definition and function in science and in law, with much misunderstanding of the concept being acknowledged amongst practitioners as well as the courts (Christensen 2014). For example, reliability of analytical methods types of error (which could be established by running the same sample multiple times to see how consistent the results is and how often an 'error' is identified) is different from errors made due to lack of poor professional standards, training, and human misconduct. This also significantly differs from unconscious errors that may arise due to cognitive biases interfering in the decision-making process.

However, in order for errors to be observable when attempting to establish validity of methods, the ground truth must be known. This is problematic in the forensic context as it is very rare that the ground truth can be known, even in cases where the evidence is corroborated by the verdict arrived at in court. Compared to experimental studies where known samples are used and there is *a priori* knowledge in order to categorically determine a correct or incorrect conclusion, in forensic cases, such knowledge is lacking. Although it has been argued that it is imperative to have different concepts of error in order to design research that will help to produce methods with known limitations, (Kloosterman et al 2014, Christensen et al. 2014, PCAS report 2016) arguably there are other factors and variables in forensic cases that are not necessarily controllable. Moreover, this becomes ever more problematic when trying to account for 'cognitive errors' that appear to occur without awareness or intention. This thesis has highlighted that there are particular variables, such as context, that can directly affect interpretations in methodologies used in forensic anthropology, which are not always within our control, and therefore, in a forensic context, can prove problematic to measure and decipher when trying to establish 'known errors' within the discipline, or within a certain method.

More importantly, there needs to be a cultural change in how we generally talk and approach the limitations within our domains. Naturally, people do not like to talk about ‘mistakes’. A common trait however, within both the lay and forensic science communities is the tendency to blame ‘bad apple’ forensic scientists for making ‘erroneous’ decisions and interpretations, and an assumption that removing such ‘bad apples’ will ensure the reliability of forensic science (Thompson 2010; Bunkley 2008). Arguably, all human decision-makers have the potential to make ‘bad apple’ decisions due to the very nature of human decision-making itself (e.g.Kassin et al. 2013; Earwaker and Nakhaeizadeh 2017). This is perhaps even more notable when using methods heavily reliant on human judgments. The results of this thesis further emphasise that it is only by acknowledging the limitations and uncertainties inherent in subjective decision-making, that the forensic community can begin to migrate towards a transparent culture, embracing a dialogue that openly explores decision-making within the forensic process; determining where issues exist, increasing understanding of the human interpretation processes involved, and finding ways in which decision-making processes can be improved.

Change needs to stem from the bottom-up, where decision-making theories are incorporated in research and practice-led teaching at an early stage. This will foster a culture of change, where future practitioners will more fully comprehend how decisions are made, and how to enhance performance in the judgment and decision-making taking place in forensic reconstructions (Dror 2016; Dror 2014b). Embedding the inclusion of decision-making as part of the forensic science process through education and training is currently lacking within the educational system in forensic anthropology, where the teaching standards are increasingly diverse. Indeed, there is still a distinct lack of clarity just how the body of knowledge concerning the application of decision theories within forensic science can be beneficial in the educational process.

The results from this thesis with regards to cognitive limitations involved in visual assessments do raise important questions with respect to current teaching and practice within the domain. As previously highlighted, there is much improvement to be made within forensic anthropological methods, with considerable development being made within traditional visual assessments and new method developments (e.g. Mahakkanukrauh et al. 2016; Konigsberg et al. 2009; Biwasaka et al. 2012). Many still however prefer to teach and practice what they have become accustomed to, and may be reluctant to diverge from such traditions.

The outcomes of this thesis should challenge some of the limitations within traditional methods that have historically been taught and used within the domain for many years. However, research is on going, and should be applied with an open mind as one cannot stress enough that the educational system needs to evolve based on the knowledge that we currently possess, not only on the collection and analysis stages of the forensic process, but also, and especially, in the interpretation and presentation phases, aiming to break down boundaries through a multidisciplinary teaching approach. Fostering change from inception will help nurture a willingness to migrate from standardized theories, and provide the tools for practitioners to incorporate unravelled logic within current practice when this becomes available through empirical research.

6.6. Future directions

Other forensic disciplines have shown and accepted that human decision-making (particularly in the difficult and ambiguous cases) is vulnerable to unconscious context effects. The discipline of forensic anthropology is not an exception. Tackling potential context effects in forensic anthropology is not an easy task due to the complexities of the decision-making involved which must often be made in line with existing policies or procedures. This thesis has begun to highlight the existence of cognitive and contextual effects through the provision of empirical data, identifying situations in which they may occur, and seeking to identify the steps that can be taken to address this issue in the future.

It is acknowledged that one limitation of the current research is that participants were not experts or working professionals within forensic anthropology. Although participants did come from appropriate backgrounds and had training within physical anthropological and osteological methods, future research should explore the role experience plays in the effect of contextual and cognitive biases. Many of the visual methods used in forensic anthropology are claimed to be subjective in nature, and generally reliant upon observation and the specialised experience of the observer (e.g. White & Folkens 2005). Although this may mean that experts within forensic anthropology are more skilled in applying certain methods and distinguishing between bone morphology and diversity, this does not necessarily mean that they are less susceptible to cognitive biases by virtue of their training and experience. Indeed empirical studies within forensic anthropology (Nakhaeizadeh et al. 2014a Klales 2016) and other forensic domains (e.g. Dror & Rosenthal 2008) have shown that contextual and confirmation bias can impact and influence the interpretations and opinions of experts in the same way as outlined in the studies presented in this thesis. Some have argued that being an expert may in fact render one increasingly prone to cognitive biases (see Dror 2011). Experts are more reliant on top down information and shortcuts (e.g. schemas) that enable experts in a given domain to quickly process information, improving efficiency in decision-making (Tversky and Khaneman 1974). This is often of great value. However, for the same aforementioned reasons, this could also create 'bias' in how 'bottom up' information is processed as it may be driven by previous experiences and expectations.

Paradoxically, as one accumulates more experience, becoming more reliant upon top-down shortcuts, with growing confidence in ability as a result of such experience and expertise, this could subconsciously lead to increasing exposure to cognitive biases. Therefore, empirical studies within the role of expertise in forensic anthropology and cognitive biases should be further explored. Although sometimes logistically difficult to implement, future studies should replicate the observed effects in the studies presented in this thesis amongst experienced professionals in order to understand not only the degree to which cognitive biases influence professionals but critically whether experience does or does not play an increasing role in the effect of cognitive biases (Dror 2011, Butt 2013).

Further, the relationships with regards to confidence in how judgments are made, and upon judgment accuracy, should be empirically tested. This thesis has indicated that there is a relationship between confidence and the interpretations of participants. However, due to the fact that there was an absence of a 'ground truth' for the majority of the skeletal remains used (except for the casts used in experimental study 3) no observations could be made as to whether examiners may be prone to overconfidence. Previous research in confidence and judgment accuracy has shown that training and experience tend to increase the confidence with which judgments are made, but does not necessarily improve the accuracy of those judgments (e.g. Jordan et al. 2013, Meissner and Kassin 2002). Further research in forensic anthropology should consider examining the relationship between judgment accuracy and confidence levels.

This thesis also focused upon contextual effects in visual methodologies used in forensic anthropology, where future studies may seek to further explore cognitive and contextual biases within metric assessments. Understanding the uncertainties inherent in both non-metric and metric methods (from a cognitive perspective) is crucial in order to assess the limitations of the technique used. Quantifying methods may result in a reduction of subjective interpretations of bones; yet, further studies need to be undertaken in order to understand the role human cognition plays in metric and technological solutions. As explained in chapter 2, a number of recent studies in forensic anthropology have demonstrated that new technology can be harnessed to develop increasingly objective metric tests in sex, ancestry and age at death. Therefore, it is important to assess whether technological solutions to cognitive biases could potentially be valuable as cognitive offloading on technology may still create further cognitive interpretation issues (Dror et al. 2012). Not only is it important to understand the practical challenges in metric methods, but also to understand the function of cognitive issues in each method applied. This should also be further examined, within, for example, trauma analyses, taphonomical processes, as well as decomposition phases, within all of which many of the interpretations are based on human judgments.

It is also important to further highlight that the contextual effects and consistency in the interpretations of participants were measured on the basis of the overall answer from all the methods available, combined for each element, rather than on one technique or the single traits scored for each method. As discussed throughout the experimental chapters, this is an important point, as previous validation and classification studies within methods used in forensic anthropology have generally shown visual methods to be reliable, with high classification accuracy, specifically for sex estimation on the pelvis (e.g. Cabo 2012). The aim of this thesis was not to conduct a validation and classification study of non-metric methods used in forensic anthropology, but rather to look further into the role context may play in the visual assessment of skeletal remains, in order to create some ‘baseline’ studies of the potential contextual effects on the interpretation of skeletal elements.

Future studies could be carried out to gain further in-depth analysis on each individual method conducted, in order to estimate whether different skeletal elements (burnt, fragmented, complete, ambiguous, non-ambiguous) and their level of ‘judgment difficulty’ play a role in the effect of context within the method used. This may include focusing on a single method (e.g. scoring system of the skull) on larger skeletal samples (multiple skulls within a specific population). As mentioned in chapter 2, studies by Klales et al. 2016 showed a strong confirmation bias amongst experts when looking at the traits and methodology outlined by Phenice (1969) on ten os coxa. Similarly, future studies should further explore the avenues of contextual and cognitive limitations within each method on a larger skeletal sample. Equally, using a known sample size, preferably from a modern population, could also aid in understanding the role of contextual effects in forensic population based studies.

The concept of cognitive bias is just one of a number of theories discussed within the field of judgment and decision-making, some of which may have the potential to enhance performance and decision efficiency. The development of a sound scientific knowledge base has also been aided by researchers who have begun to look beyond cognitive bias within forensic science, taking into consideration the wider applications of judgment and decision-making theories (e.g. Biedermann et al. 2016; Gittelson et al. 2013; Englich & Musseweiler 2001) Together, these studies alongside the pioneering studies of several scholars within cognitive biases and human interpretation issues in forensic science and criminal investigations (e.g. Kassin et al 2013; Dror & Charlton 2006; Miller 1984; Thompson; 2009 Saks et al. 2003) has produced a valuable insight into the central role of cognition in forensic science, creating a rise in what may be referred to as cognitive forensics. Cognitive forensics goes beyond the issues of confirmation and contextual biases, and deals with all forms of judgment and decision-making involved in forensic reconstruction. Therefore, future studies within forensic anthropology may need to explore the broader aspect of decision theories to fully comprehend not only how examiners reach conclusions, but also how research in cognition could enhance forensic anthropological practice and procedures.

For example, the application of ‘priming’ as a tool to better understand decisions, including those made in fingerprint analysis (see Earwaker et al. 2017a) could be applied within forensic anthropology. The nature of analyses within some identification fields (comparison and classification, and pattern recognition respectively), allow for a relatively straightforward application of methods and transferable solutions into other forensic disciplines of a similar nature.

Furthermore, there is a requirement for an increasingly structured, exhaustive, inclusive, and sustainable approach to improve decision-making within not only forensic anthropology, but also forensic science in general. This could arguably be achieved through a holistic approach, taking into consideration the interactions and impact of a decision process within a wider context of interactions within the forensic science process (crime scene to court) and stakeholders in the criminal justice system. Although the research in this thesis looked into both the early stage of skeletal collection as well as later stage of skeletal interpretation, many of the studies conducted in cognitive biases (and judgment and decision-making in general) have predominantly focused upon the interpretation phase within a forensic science process (Earwaker et al. 2017b). Only by conducting research that allows for a more structured and transparent examination of the decisions and their interdependencies that are made throughout the forensic process, as well as a greater understanding of the wider decision ecology of the criminal justice system, can one improve the quality, reliability, and efficiency of forensic interpretation.

Chapter 7. Conclusion

7.1. Summary of the thesis aim and research questions

As highlighted in Chapter 1, the aim of the present thesis was to further examine the extent to which cognitive biases are present within forensic anthropological methods. More specifically, the thesis sought to understand the degree of contextual effects in forensic anthropological assessments, and thereby identify the means to avoid potential cognitive biases that might arise from interpretation issues. The research conducted was concerned with addressing the knowledge gap in forensic anthropology with regards to contextual effects in biological profiling, focusing on three main research questions:

1. Does contextual information affect previous judgments when assessing skeletal remains of an ambiguous nature?

2. Does early exposure to 'extraneous' contexts in the excavation of skeletal remains cascade and thereby affect the subsequent assessment of the skeletal analysis?

3. Does the order of examination of skeletal remains a) influence the interpretation of the subsequent skeletal element, and b) act as an influence and determine the final conclusion of the assessment?

A series of experiments were designed to test for cognitive and contextual effects empirically within forensic anthropological methods and procedures. A holistic examination of the stages and methodological procedures was undertaken to establish when, and to what extent cognitive factors may affect performances and render the judgements of participants to be compromised, and equally when they do not.

7.2. Research question 1. Does contextual information affect previous judgments when assessing skeletal remains of an ambiguous nature?

Experimental studies 1 and 2 presented in chapter 3 specifically address whether contextual information can affect previous judgments of the participants when assessing skeletal remains of an ambiguous nature. The experiments were carried out in order to gain insights into whether the decisions of participants were independent, and thus consistent regardless of contextual influences, or alternatively, whether the participants changed their previous decisions as a result of being given context.

The results of the experimental findings that were presented and discussed in Chapter 3 and subsequently discussed in Chapter 6 unequivocally identified that additional grave context descriptions and artifacts, as well as osteological reports, could have an influence on the judgments of participants in the visual assessments of the skeletal remains. More importantly, the results from experiment 1 and 2 not only showed that visual assessments are mediated by context, but also that context can change and override participants previous assessments on the same skeletal remains.

The result from the study has demonstrated empirically that context could potentially have a powerful role in forensic anthropological decision-making, especially in conditions of ambiguity. For example, across all the visual methods addressed in experiment 2, 62.8% of the participants were affected by the contextual information (changing their decisions to align with the context provided), showing a lack of consistency in the interpretation of the skeletal remains when context is given. This further highlighted that the vulnerability to contextual influences was found across all traditional visual methods (sex, ancestry, and age at death), offering insight into factors that may mediate human performance. The results from chapter 3 have contributed data in achieving the overall aim of this thesis, and acknowledged the need for further research within contextual biases in biological profiling. This would mean looking further into the consistency of expert performance, and how this relates to potential 'biasability' within and between experts when context is given in the assessment of known skeletal remains.

7.3. Research question 2. Does early exposure to 'extraneous' contexts in the excavation of skeletal remains cascade and thereby affect the subsequent assessment of the skeletal analysis?

As outlined in Chapter 4, experimental study 3 investigated the potential effects of initial exposure to context at a crime scene upon judgment and decision-making. The study specifically addressed whether clothing associated with skeletal excavations could influence and impact the evaluations and judgments of the participants. This was to be done in order to examine whether early exposure to such contexts would cascade and affect the subsequent assessment of the skeletal remains.

The findings from experimental 3 showed that ‘extraneous’ grave clothing associated with skeletal excavations impacts upon initial beliefs, interpretations, confidence level of judgments, and the subsequent assessments of skeletal remains. Unexpectedly, experiment 3 (Chapter 4) showed that early exposure of a strong indication of a certain outcome did in fact ‘bias’ the interpretation of non-ambiguous male skeletal casts. In contrast to previous research (e.g. Nakhaeizadeh, Itiel E Dror, et al. 2014; Nakhaeizadeh, Hanson, et al. 2014; Dror et al. 2011; Dror et al 2006), the results from experimental study 3 showed that contextual effects can also have an impact on non-ambiguous evidence. For example, all participants in the control group and majority of participants in the ambiguous context group (75%) estimated the skeletal cast to be male. However only 1 participant in the female context group estimated the same skeletal cast to be of male. The results showed a bias cascade effect in which bias arises as a result of extraneous information, cascading from one stage to another (e.g. skeletal excavation to skeletal analysis), which could conceivably, severally impinge the scientist accurately interpreting evidence. In addition, the results concluded that not all contexts in a given situation would affect the final assessment, showing (compared to experiment 1 and 2) that when dealing with non-ambiguous skeletal remains, only a very ‘strong’ context altered the decision-making process of the participants.

The result from experimental study 3 contributes to the current debate with regards to how to control the information flow between the different stages of the forensic investigation. The data provides an important step towards understanding the potential effects of initial exposure to contextual effects at a crime scene upon judgment and decision-making within forensic anthropology (as well as the broader forensic science domain), which has previously not been established empirically. Indeed, the results have highlighted the potential risk of the anthropologist being exposed to ‘extraneous information’ at a very early stage of a forensic investigation with future work needed in order to fully understand when contextual biases may cascade in forensic investigations. Whilst not all contexts will have a biasing influence, experimental study 3 from this thesis does illustrate that it is possible for context to affect early hypothesis and decision-making, even with non-ambiguous evidence. It is therefore important to be aware of such instances, and to take steps to ensure that inferences are shielded as much as possible form potential cascaded bias from the exposure of context.

7.4. Research question 3. Does the order of examination of skeletal remains a) influence the interpretation of the subsequent skeletal element, and b) act as an influence and determine the final conclusion of the assessment?

Chapter 5, experimental study 4 focused on a different type of influence that may arise due to certain procedural practice. The study sought to establish whether the order in which the skeletal elements were assessed had an affect on the interpretations of the following skeletal element, as well as the overall assessment. The results showed no such influence when dealing with non-ambiguous traits. The results concluded that the sex estimation on the skull and the os coxa was not influenced depending on the order in which the skeletal elements were assessed. In addition, the order of examination did not influence the final and combined interpretation of the skeletal remains.

The data did however provide some valuable preliminary insight into the tendency for some of the participants (40%) to rely on the precise sex of the os coxa, rather than the precise sex of the skull when making a definitive overall sex estimation. The results raise some potential limitations concerning the value of the skull as a sex indicator for forensic cases, especially in events where the skull is the only skeletal element present, or the innominate being badly damaged, as well as when dealing with commingled remains (e.g. mass grave excavations and disaster victims cases). Although the data from experimental study 4 did not show any 'cognitive influences' arising due to procedural practice and order of examination, the results still provide some insight into when certain factors may affect performance, and render the judgements of participants compromised, and equally when they do not. More importantly, this further raised some concerns and possibilities for future research with regards to potential issues that may arise in the decision-making process of comingled skeletal remains.

7.5. Contribution to the body of knowledge within Forensic anthropology and Forensic science

A better understanding of the underlying processes of the decisions being made and the extent to which contextual influences occur in forensic anthropology needs to be acknowledged and addressed. The main findings from chapter 3, 4, and 5 of this thesis conclude that context is influential in visual methods used in forensic anthropology (in the establishment of a biological profile) where human judgment plays a central role. In addition, the result shows that the power of contextual influences in the assessment of skeletal remains will differ, and in some cases may result in contextual biases. Although the methods used in visual assessments in forensic anthropology are considered 'foundationally valid' and in principle reliable, there are still many reasons why examiners may not always reach robust conclusions. The results from this research, (specifically chapter 3 and 4) has shown that one of these reasons could be the deficiency in acknowledging the role and extent of cognitive and contextual influences within the discipline. Thus, it is important for the forensic anthropological community not to underestimate and minimise the importance of these issues as it has not only shown to affect expert interpretations across numerous forensic disciplines, (e.g. Dror et al. 2006; Dror & Hampikian 2011; Osborne et al. 2014; Stoel et al. 2014; Nakhaeizadeh et al. 2014; Page et al. 2012; Archer & Wallman 2016) but to also affect the human role at the different stages of the forensic science process.

More importantly, the results from this thesis have highlighted that what is considered as influential will depend on:

- the nature of the task,
- the ambiguity level of given characteristics being interpreted,
- the difficulty of the judgment,
- and the strength of the context in which the decision is made.

Therefore, the data produced from this thesis has added to the existing body of knowledge, aiding and improving our understanding of human decision-making within the forensic anthropological process, which arguably may have previously been disregarded.

Furthermore, cognitive and contextual influences have for too long been misunderstood and neglected within the forensic domains. Considering how science and law continue to interrelate, and that the issue of scientific standards within the forensic disciplines is under scrutiny, (as highlighted throughout Chapter 2) the forensic science community must be committed to not only continuing to address the issue of cognitive and contextual biases, but also to ensure the most effective implementation of valid solutions. This research has highlighted the need for improvement in our understanding of human decision-making not only within forensic anthropology but the wider forensic science disciplines, calling for a more holistic, comprehensive, and transparent examination of each decision individually, and interdependently within the commonly considered linear forensic science process. The data collected throughout this thesis has highlighted that there are particular variables, such as context, that can directly affect interpretations in methodologies used where human interpretation plays a central role. These variables affecting our interpretations are not always within our control and can, in a forensic context, prove to be problematic to measure and decipher. Therefore, the results from the experimental studies in this thesis have further acknowledged that it will be important to offer approaches that can not only be sufficiently generalised across different investigations and sample examination, but that are also sufficiently context sensitive to each case and each sample within it. It is only by acknowledging the limitations and uncertainties inherent in subjective decision-making, that the forensic community can begin to migrate towards a transparent culture, embracing a dialogue that openly explores decision-making within the forensic process; determining where issues exist, increasing understanding of the human interpretation processes involved, and finding ways in which decision-making processes can be improved.

7.6. Avenues for future research

The findings of this research have highlighted a number of areas in which further research is warranted. With regards to contextual biases in forensic anthropology, further work should be undertaken exploring the avenues of contextual effects within each separate method (metric and non-metric) on a larger skeletal sample. In a manner that complements the findings of this thesis, experimental research efforts that further investigate contextual effects using a known sample size, preferably from a modern population, could aid in understanding the role of context in forensic population based studies. Conducting studies on a known sample size will allow for further studies measuring the relationship between cognitive biases and judgment accuracy. Further, the relationships with regards to confidence in how judgments are made, and upon judgment accuracy, should be empirically tested. This thesis has indicated that there is a relationship between confidence and the interpretations of participants; however, concerted attempts to understand how this relates to overconfidence in judgments could reveal further evidential value on possible misleading interpretations of skeletal remains.

Additionally, possible studies could explore the impact of context within other aspects of forensic anthropology, such as trauma analysis, pathology, as well as burnt skeletal remains, with many of the methods currently used within these areas of forensic anthropology heavily based on visual interpretations, training, and expertise. Therefore, empirical studies within the role of expertise in forensic anthropology and cognitive biases should be further explored. As discussed in Chapter 6, although sometimes logistically difficult to implement, future studies should replicate the observed effects in the studies presented in this thesis amongst experienced professionals in order to understand not only the degree to which cognitive biases influence professionals, but critically, whether experience does or does not play an increasing role in the effect of cognitive biases. The study of expertise and cognitive biases should also consider testing for other variables such as emotions and time pressure in order to see whether different stimuli affect performance and interpretations of skeletal remains. This is important, as preliminary studies within the domain have shown that high emotional context could result in expectation bias when assessing trauma on skeletal remains (Nakhaeizadeh, Hanson, et al. 2014).

As discussed in Chapter 6, the concept of cognitive bias is just one of a number of theories discussed within the field of judgment and decision-making, some of which may have the potential to enhance performance and decision efficiency. The concept of cognitive forensics (discussed in Chapter 6) goes beyond the issues of confirmation and contextual biases, and deals with all forms of judgment and decision-making involved in forensic reconstruction. Therefore, future studies within forensic anthropology may need to explore the broader aspect of decision theories to fully comprehend not only how examiners reach conclusions, but also how research in cognition could enhance forensic anthropological practice and procedures. For example, the nature of analyses within some identification fields (comparison and classification, and pattern recognition respectively), allow for a relatively straightforward application of methods and transferable solutions into other forensic disciplines of a similar nature. The forensic anthropology community would benefit in exploring the wider aspect of cognitive bias in judgment and decision-making within other fields, where suggested solutions may be transferable.

In addition, furthering our understanding of the role that cognition plays in the collection, analysis, interpretation and presentation of evidence, will enable the forensic anthropological community (and the forensic sciences broadly) to holistically approach cognitive biases in addition to developing methods and approaches that will aim to move away from subjective interpretations. This approach will aid in addressing some of the concerns raised by the NAS (2009) report, as well as the PCAST (2016) report with regards to interpretations issues.

In conclusion, the forensic anthropological community has come far in the development of the discipline. The findings of this thesis highlight that just as other disciplines have started to act and entered into a dialogue of cognitive interpretation issues, further research within the forensic anthropology domain is needed to articulate and develop frameworks that incorporate an understanding of when context may influence the interpretation of expert evidence. Addressing cognitive issues for interpretation of evidence will be crucial for forensic anthropology to remain as a valuable approach within forensic investigations. Although the results from this thesis has only begun to unravel some of the contextual effects in forensic anthropology, recognising the limitations of current methods and identifying weaknesses, including cognitive biases, will enable further developments, and allow forensic anthropologists to offer increasingly robust, accurate, and valuable intelligence and evidence to investigations.

References

Aarts, H. & Dijksterhuis, A., 2002. Category activation effects in judgment and behaviour: The moderating role of perceived comparability. *British Journal of Social Psychology*, 41, pp.123–138.

Anders Ericsson, K. et al., 1993. The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review*, 100(3), pp.363–406.

Anderson, C.A. & Kellam, K.L., 1992. Belief perseverance, biased assimilation, and covariation detection: The effect of hypothetical social theories and new data. *Personality and Social Psychology Bulletin*, 18, pp.555–565.

Anderson, C.A., Lepper, M.R. & Ross, L., 1980. Perseverance of Social Theories: The Role of Explanation in the Persistence of Discredited Information. *Journal of Personality and Social Psychology*, 39(6), pp.1037–1049.

Anderson, J.R., 2000. *Cognitive Psychology and Its Implications Fifth 5th Edition* 5th ed., United State of America: Worth Publishers and WH Freeman.

Appleby, S.C. and Kassin, S.M., 2016. When self-report trumps science: Effects of confessions, DNA, and prosecutorial theories on perceptions of guilt. *Psychology, Public Policy, and Law*, *22*(2), p.127.

Archer, M.S. & Wallman, J.F., 2016. Context Effects in Forensic Entomology and Use of Sequential Unmasking in Casework. *Journal of forensic sciences*, 61(5), pp.1270-1277

Asala, S.A. et al., 2001. Sex determination from the head of the femur of South African whites and blacks. *Forensic science international*, 117(1–2), pp.15–22.

Asch, S.., 1951. Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow, ed. *Groups, leadership and men; research in human relations*. Carnegie Press, pp. 177–190.

Ask, K. & Granhag, P.A., 2005. Motivational sources of confirmation bias in criminal investigations: the need for cognitive closure. *Journal of Investigative Psychology and Offender Profiling*, 2(1), pp.43–63.

Baccino, E. & Schmitt, A., 2006. Determination of Adult Age at Death in the Forensic Context. In *Forensic Anthropology and Medicine*. Totowa, NJ: Humana Press, pp. 259–280.

Balcetis, E. & Dunning, D., 2006. See what you want to see: motivational influences on visual perception. *Journal of personality and social psychology*, 91(4), pp.612–25.

Bandura & Albert, 1986. *Social foundations of thought and action: A social cognitive theory.*, Prentice-Hall, Inc.

Baron, J., 1995. Myside bias in thinking about abortion. *Thinking and Reasoning*, 1, pp.221–235.

Bergman, O. et al., 2010. Anchoring and cognitive ability. *Economics Letters*, 107(1), pp.66–68.

Bhalla, M. & Proffitt, D.R., 1999. Visual-motor recalibration in geographical slant perception. *Journal of experimental psychology. Human perception and performance*, 25(4), pp.1076–96.

Bidmos, M.A., Gibbon, V.E. & Štrkalj, G., 2010. Recent advances in sex identification of human skeletal remains in South Africa. *South African Journal of Science*, 106(11/12), p.6.

Bieber, P., 2012. Measuring the impact of cognitive bias in fire investigation. International symposium on fire investigation. *Science and Technology*, pp.3–15.

Biedermann, A., Bozza, S. & Taroni, F., 2016. The decisionalization of individualization. *Forensic Science International*, 266, pp.29–38.

Biwasaka, H. et al., 2009. Analyses of sexual dimorphism of contemporary Japanese using reconstructed three-dimensional CT images--curvature of the best-fit circle of the greater sciatic notch. *Legal medicine (Tokyo, Japan)*, 11 (1), pp.260-2.

Biwasaka, H. et al., 2012. Analyses of sexual dimorphism of reconstructed pelvic computed tomography images of contemporary Japanese using curvature of the greater sciatic notch, pubic arch and greater pelvis. *Forensic Science International*, 219(1–3), p.288.e1-288.e8.

Blackwell, S.E. & Holmes, E.A., 2010. Modifying interpretation and imagination in clinical depression: A single case series using cognitive bias modification. *Applied Cognitive Psychology*, 24(3), pp.338–350.

Blau, S., 2016. How traumatic: a review of the role of the forensic anthropologist in the examination and interpretation of skeletal trauma. *Australian Journal of Forensic Sciences*, pp.1–20.

Bressan, P. & Dal Martello, M.F., 2002. Talis pater, Talis filius: perceived resemblance and the belief in genetic relatedness. *Psychological Science*, 13, pp.213–218.

Bright, D.A. & Goodman-Delahunty, J., 2006. Gruesome Evidence and Emotion: Anger, Blame, and Jury Decision-Making. *Psychiatry, Psychology & Law*, 11(1), pp.154–166.

Brooks, S. & Suchey, J.M., 1990. Skeletal age determination based on the os pubis: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Human Evolution*, 5(3), pp.227–238.

Bruner, J.S. & Potter, M.C., 1964. Interference in visual recognition. *Science (New York, N.Y.)*, 144(3617), pp.424–5.

Bruzek, J., 2002. A method for visual determination of sex, using the human hip bone. *American journal of physical anthropology*, 117(2), pp.157–68.

Buckberry, J.L. & Chamberlain, A.T., 2002. Age estimation from the auricular surface of the ilium: A revised method. *American Journal of Physical Anthropology*, 119(3), pp.231–239.

Budowle, B. et al., 2009. A perspective on errors, bias, and interpretation in the forensic sciences and direction for continuing advancement. *Journal of forensic sciences*, 54(4), pp.798–809.

Bugelski, B.R. & Alampay, D.A., 1961. The role of frequency in developing perceptual sets. *Canadian Journal of Psychology Revue Canadienne de Psychologie*, 15(4), pp.205–211.

Bukistra, J.E. & Ubelaker, D.H., 1994. Documentation of sex differences and age changes in adults. In *Standards for data collection from human skeletal remains*. Fayetteville, AR: Arkansas: Archaeological Survey Research Series No. 44, pp. 15–38.

Bunkley, N., 2008. Detroit Police Lab Is Closed After Audit Finds Serious Errors In Many Cases. *The New York Times*.

Burke, A.S., 2005. Improving Prosecutorial Decision Making: Some Lessons of Cognitive Science. *William and Mary Law Review*, 47(5), pp.1587–1633.

Burns, K.R., 1999. *Forensic anthropology training manual*, Prentice Hall.

Butt, L., 2013. The forensic confirmation bias: Problems, perspectives, and proposed solutions—Commentary by a forensic examiner. *Journal of Applied Research in Memory and Cognition*, 2(1), pp.59–60.

Byers, S.N., 2010. *Introduction to Forensic Anthropology (4th Edition)* S. N. Byers, ed., Pearson Education.

Byrne, A. & Eysenck, M.W., 1993. Individual differences in positive and negative interpretive biases. *Personality and Individual Differences*, 14(6), pp.849–851.

Cabo, L.L., Brewster, C.P. & Luengo Azpiazu, J., 2012. Sexual dimorphism: interpreting sex markers. In D. C. Dirkmaat, ed. *A Companion to Forensic Anthropology*. Chichester, UK: Blackwell Publishing, Ltd, pp. 248–286.

Caddy, B. & Cobb, P., 2009. Forensic Science. In White P.C, ed. *Crime Scene to Court: The essentials of Forensic Science. 2nd ed.* Camebridge: Royal Society of Chemistry.

Cattaneo, C. et al., 1999. Determining the human origin of fragments of burnt bone: a comparative study of histological, immunological and DNA techniques. *Forensic Science International*, 102(2–3), pp.181–191.

Cattaneo, C., 2007. Forensic anthropology: developments of a classical discipline in the new millennium. *Forensic science international*, 165(2–3), pp.185–93.

Caverni, J.P., Fabre, J.-M. & Gonzalez, M., 1990. Cognitive Biases: Their Contribution for Understanding Human Cognitive Processes. *Advances in Psychology*, 68, pp.7–12.

Champod, C., 2014. Research focused mainly on bias will paralyse forensic science. *Science & justice : journal of the Forensic Science Society*, 54(2), pp.107–9.

Champod, C., Evett, I.W. & Kuchler, B., 2001. Earmarks as evidence: a critical review.

Journal of forensic sciences, 46(6), pp.1275–1284.

Charlton, D., Fraser-Mackenzie, P.A.F. & Dror, I.E., 2010. Emotional experiences and motivating factors associated with fingerprint analysis. *Journal of forensic sciences*, 55(2), pp.385–93.

Charman, S.D. & Wells, G.L., 2008. Can eyewitnesses correct for external influences on their lineup identifications? The actual/counterfactual assessment paradigm. *Journal of experimental psychology. Applied*, 14(1), pp.5–20.

Cheetham, P. & Hanson, I., 2009. Excavation and Recovery in Forensic Archeological Investigations. In Soren Blaue and Douglas H Ubelaker, ed. *Handbook of Forensic Anthropology and Archeology*. Left Coast Press Inc, pp. 141–152.

Chen, X.-P. et al., 2011. Determining the age at death of females in the Chinese Han population: using quantitative variables and statistical analysis from pubic bones. *Forensic science international*, 210(1–3), p.278.

Cheng, P.W. et al., 1986. Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology*, 18(3), pp.293–328.

Christensen, A.M. et al., 2014. Error and its meaning in forensic science. *Journal of forensic sciences*, 59(1), pp.123–6.

Christensen, A.M., 2004. The impact of Daubert: implications for testimony and research in forensic anthropology (and the use of frontal sinuses in personal identification). *Journal of forensic sciences*, 49(3), pp.427–30.

Christensen, A.M. & Crowder, C.M., 2009. Evidentiary standards for forensic anthropology. *Journal of forensic sciences*, 54(6), pp.1211–6.

Clark, M.A. et al., 2015. Quantification of Maxillary Dental Arcade Curvature and the Estimation of Biological Ancestry in Forensic Anthropology. *Journal of forensic sciences*.

Cooley, C. & Turvey, B.E., 2011. Observer Effect and Examiner Bias: Psychological Influences on the Forensic Examiner. In J. W. Chisum & B. . Turvey, eds. *Crime Reconstruction*. Elsevier, pp. 61–90.

Crispino, F., Ribaux, O. & Margot, P., 2011. Forensic science – A true science? *Australian Journal of Forensic Sciences*, 43(2–3).

Crowder, C. & Stout, S.D., 2011. *Bone histology : an anthropological perspective*, CRC Press.

Cunha, E. & Cattaneo, C., 2006. Forensic Anthropology and Forensic Pathology. In *Forensic Anthropology and Medicine*. Totowa, NJ: Humana Press, pp. 39–53.

Dirkmaat, D.C. et al., 2008. New perspectives in forensic anthropology. *American journal of physical anthropology*, Suppl 47(51), pp.33–52.

Dirkmaat, D.C. & Cabo, L.L., 2012. Forensic Anthropology:Embracing the New Paradigm. In D. C. Dirkmaat, ed. *A companion to forensic anthropology*. Blackwell Publishing Ltd, pp. 3–40.

Dominguez, V.M. & Crowder, C.M., 2012. The utility of osteon shape and circularity for differentiating human and non-human Haversian bone. *American journal of physical anthropology*, 149(1), pp.84–91.

Dror, I.E., 2009. How can Francis Bacon help forensic science? The four idols of human biases. *Jurimetrics: The Journal of Law, Science, and Technology*, 50, pp.93–110.

Dror, I.E., 2011. The Paradox of Human expertise: why experts get it wrong. In N. Kapur, ed. *The Paradoxical Brain*. Cambridge, pp. 177–188.

Dror, I.E., 2012. Cognitive Bias in Forensic Sciense. In *The 2012 yearbook of science & technology*. New York: McGraw-Hill, pp. 43–45.

Dror, I.E., 2014. Practical Solutions to Cognitive and Human Factor Challenges in Forensic Science. *Forensic Science Policy & Management: An International Journal*, 4(3–4), pp.105–113.

Dror, I.E., 2015. Cognitive neuroscience in forensic science: understanding and utulizing the human element. *Philosophical Transactions of the Royal Sociaty*, 370(1674).

Dror, I.E., 2016. A Hierarchy of Expert Performance (HEP). *Journal of Applied Research in Memory and Cognition*, 5, pp.121–127.

Dror, I.E. & Kosslyn, S.M., 1998. Age degradation in top-down processing: Identifying objects from canonical and noncanonical viewpoints. *Experimental Aging Research*, 24(3), pp.203–216.

Dror, I.E., et al., 2005. When emotions get the better of us: the effect of contextual top-down processing on matching fingerprints. *Applied Cognitive Psychology*, 19(6), pp.799–809.

Dror, I.E., & Charlton, D., 2006. Why Experts Make Errors. *Journal of Forensic Identification*, 56(4), pp.600–616.

Dror, I.E., Charlton, D. & Péron, A.E., 2006. Contextual information renders experts vulnerable to making erroneous identifications. *Forensic science international*, 156(1), pp.74–78.

Dror, I.E., & Fraser-Mackenzie, P.A.., 2008. Cognitive Biases in Human Perception, Judgment and Decision Making: Bridging Theory and the Real World. In K. Rossmo, ed. *Criminal Investigative Failures*. Taylor and Francis CRC Press, pp. 53–57.

Dror, I.E., & Rosenthal, R., 2008. Meta-analytically quantifying the reliability and biasability of forensic experts. *Journal of forensic sciences*, 53(4), pp.900–3.

Dror, I.E. & Cole, S.A., 2010. The vision in "blind" justice: Expert perception, judgment, and visual cognition in forensic pattern recognition. *Psychonomic bulletin & review*, 17(2), pp.161–167.

Dror, I.E., & Hampikian, G., 2011. Subjectivity and bias in forensic DNA mixture interpretation. *Science and Justice*, 51(4), pp.204–208.

Dror, I.E., et al., 2011. Cognitive issues in fingerprint analysis: inter- and intra-expert consistency and the effect of a "target" comparison. *Forensic science international*, 208(1–3), pp.10–7.

Dror, I.E., et al., 2012. The impact of human-technology cooperation and distributed cognition in forensic science: biasing effects of AFIS contextual information on human experts. *Journal of forensic sciences*, 57(2), pp.343–52.

Dror, I.E., et al., 2015. Letter to the Editor- Context Management Toolbox: A Linear Sequential Unmasking (LSU) Approach for Minimizing Cognitive Bias in Forensic

Decision Making. *Journal of Forensic Sciences*, 60, pp.1–2.

Đurić, M. et al., 2005. The reliability of sex determination of skeletons from forensic context in the Balkans. , 147.

Eagly, A.H. et al., 1999. The Impact of Attitudes on Memory: An Affair to Remember. *Psychological Bulletin*, 125(1), pp.64–89.

Earwaker, H., Charlton, D. & Bleay, S., 2015. *Fingerprinting – the UK Landscape: Processes, Stakeholders and Interactions*,

Earwaker, H., Nakhaeizadeh, S. Smit, N., & Morgan., R.M., 2017a. A cultural change to enable improved decision-making in the forensic sciences: a six phased approach *Science and Justice* (in preparation).

Earwaker, H., Nakhaeizadeh, S. Smit, N., Morgan., R.M., & Harris., A., 2017b. The role of psychological theories in the enhancement of the forensic science decision-making: An experimental study into the effect of priming *Forensic Science International* (in preperation).

Edmond, G. et al., 2014. Contextual bias and cross-contamination in the forensic sciences: the corrosive implications for investigations, plea bargains, trials and appeals. *Law, Probability and Risk*, 14(1), pp.1–25.

Edmond, G. et al., 2016. Thinking forensics: Cognitive science for forensic practitioners. *Science & Justice*, 0(0), pp.353–363.

van den Eeden, C.A.J., de Poot, C.J. & van Koppen, P.J., 2016. Forensic expectations: Investigating a crime scene with prior information. *Science & Justice*.

Effros, B. et al., 2000. Skeletal sex and gender in Merovingian mortuary archaeology. *Antiquity*, 74(285), pp.632–639.

Elliott, M. & Collard, M., 2009. Fordisc and the determination of ancestry from cranial measurements. *Biology Letters*, 5(6).

Elstein, A.S., 1999. Heuristics and biases:selected errors in clinical reasoning. *Academic Medicine Journal of the Association of American Medical Colleges*, 74(7), pp.791–794.

Englich, B. & Musseweiler, T., 2001. Sentencing Under Uncertainty: Anchoring Effects in the Courtroom. *Journal of Applied Social Psychology*, 31(7), pp.1535–1551.

Ericsson, K.A., Chamess, N. & Charness, N., 1994. Expert performance: Its structure and acquisition. *American Psychologist*, 49(8), pp.725–747.

Evans, J.S.B.T. & Pollard, P., 1990. Belief Bias and Problem Complexity in Deductive Reasoning. *Advances in Psychology*, 68, pp.131–154.

Falys, C.G. & Prangle, D., 2015. Estimating age of mature adults from the degeneration of the sternal end of the clavicle. *American Journal of Physical Anthropology*, 156(2), pp.203–214.

FBI & Federal Bureau of Investigation, 2015. FBI — FBI Testimony on Microscopic Hair Analysis Contained Errors in at Least 90 Percent of Cases in Ongoing Review. Available at: https://www.fbi.gov/news/pressrel/press-releases/fbi-testimony-on-microscopic-hair-analysis-contained-errors-in-at-least-90-percent-of-cases-in-ongoing-review [Accessed April 17, 2016].

Fenton, N. et al., 2014. When "neutral" evidence still has probative value (with implications from the Barry George Case). *Science & justice : Journal of the Forensic Science Society*, 54(4), pp.274–87.

Fenton, N., Neil, M. & Lagnado, D.A., 2012. A general structure for legal arguments about evidence using Bayesian networks. *Cognitive science*, 37(1), pp.61–102.

Fiddick, L., Cosmides, L. & Tooby, J., 2000. No interpretation without representation: the role of domain-specific representations and inferences in the Wason selection task. *Cognition*, 77(1), pp.1–79.

Findley, K.A. & Scott, M.S., 2006. The Multiple Dimensions of Tunnel Vision in Criminal Cases. *Wisconsin Law Revieve*, 2, pp.292–397.

Fingerprint Inquiry, 2011. *The Fingerprint Inquiry Scotland*,

Fischhoff, B., Slovic, P. & Lichtenstein, S., 1977. Knowing with Certainty: The Appropriateness of Extreme Confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), pp.552–564.

Forensic Science Regulator, 2016. *Codes of Practice and Conduct for Forensic Science providers and practitioners in the Criminal Justice System*,

Found, B., 2014. Deciphering the human condition: the rise of cognitive forensics. *Australian Journal of Forensic Sciences*, (October), pp.1–16.

Frank, M.G., Yarbrough, J.D. & Ekman, P., 2006. Investigative interviewing and the detection of deception. In T. Williamson, ed. *Investigative Interviewing Rights, research and regulation*. Willan Publishing, pp. 229–255.

Franklin, D. et al., 2008. Discriminant function sexing of the mandible of Indigenous South Africans. *Forensic Science International*, 179(1), p.84.e1-84.e5.

Franklin, D., 2010. Forensic age estimation in human skeletal remains: current concepts and future directions. *Legal medicine (Tokyo, Japan)*, 12(1), pp.1–7.

Fraser-Mackenzie, P.A.., Bucht, R.E. & Dror, I.E., 2013. Forensic Judgment and Decision-making. In P. H. Crowley & T. R. Zentall, eds. *Comparative Decision Making*. Oxford University Press, USA, pp. 385–415.

Fraser-Mackenzie, P.A.F., Dror, I.E. & Wertheim, K., 2013. Cognitive and contextual influences in determination of latent fingerprint suitability for identification judgments. *Science & justice : journal of the Forensic Science Society*, 53(2), pp.144–53.

Garrett, B.L. et al., 2009. Invalid forensic testimony and wrongful convictions. INVALID *Virginia Law Review*, 95(1), pp.1–97.

Garrett, B.L., 2008. Judging Innocence. *Columbia Law Review*, 108, pp.55–142.

Garvin, H.M., 2012. Adult Sex Determination. In D. C. Dirkmaat, ed. *A companion to forensic anthropology*. Blackwell Publishing Ltd.

Gianelli, P.C., 2007. Confirmation Bias. *Criminal Justice*, 22(3), pp.60–61.

Gigerenzer, G. & Gaissmaier, W., 2011. Heuristic decision making. *Annual review of psychology*, 62, pp.451–82.

Giles, E. & Elliot, O., 1963. Sex determination by discriminant function analysis of crania. *American Journal of Physical Anthropology*, 21(1), pp.53–68.

Girotto, V. & Politzer, G., 1990. Conversational and World Knowledge Constraints on Deductive Reasoning. *Advances in Psychology*, 68, pp.87–107.

Gittelson, S. et al., 2013. *Decision-theoretic reflections on processing a fingermark*,

Government Chief Scientific Adviser, 2015. *Forensic Science and Beyond: Authenticity Provenance and Assurance*,

Government Chief Scientific Advisior, 2015. *Forensic Science and Beyond: Authenticity, Provenance and Assurance, Evidence and Case Studies*,

Green, R.L., 1992. *Human Memory: Paradigms and Paradoxes - Robert L. Greene - Google Books*, New York: Lawrence Erlbaum Associates.

Grivas, C.R. & Komar, D.A., 2008. Kumho, Daubert, and the nature of scientific inquiry: implications for forensic anthropology. *Journal of forensic sciences*, 53(4), pp.771–6.

Groen, W.J.M., Márquez-Grant, N. & Janaway, R., 2015. *Forensic Archaeology : a Global Perspective.*, Wiley.

de Gruijter, M., de Poot, C. & Elffers, H., 2016. Reconstructing with trace information: Does rapid identification information lead to better crime reconstructions? *Journal of Investigative Psychology and Offender Profiling*.

Guyomarc'h, P. et al., 2011. Accuracy and reliability in sex determination from skulls: a comparison of Fordisc® 3.0 and the discriminant function analysis. *Forensic science international*, 208(1–3), p.180.e1-6.

Haglund, W.D., 2001. Archeology and Forensic Death Investigations. *Historical Archeology*, 35(1), pp.26–34.

Hamilton, D.L. & Zanna, M.P., 1974. Context effects in impression formation: Changes in connotative meaning. *Journal of Personality and Social Psychology*, 29(5), pp.649–654.

Hartnett, K.M., 2010. Analysis of Age-at-Death Estimation Using Data from a New, Modern Autopsy Sample-Part I: Pubic Bone. *Journal of Forensic Sciences*, 55(5), pp.1145–1151.

Hawkins, S.A. & Hastie, R., 1990. Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, 107(3), pp.311–327.

Hefner, J.T., 2009. Cranial nonmetric variation and estimating ancestry*. *Journal of forensic sciences*, 54(5), pp.985–95.

Hefner, J.T. et al., 2015. Morphoscopic Trait Expression in "Hispanic" Populations. *Journal of forensic sciences*.

Hefner, J.T. et al., 2007. *The value of experience, education and methods in ancestry prediction*, San Antonio Procceedings of the 59th Annual Meeting of the American Academy of Forensic Science.

Hefner, J.T. & Ousley, S.D., 2014. Statistical Classification Methods for Estimating Ancestry Using Morphoscopic Traits,. *Journal of Forensic Sciences*, 59(4), pp.883–890.

Hefner, J.T., Ousley, S.D. & Dirkmaat, D.C., 2012. Morphoscopic traits and assessment of ancestry. In D. C. Dirkmaat, ed. *A Companion to Forensic Anthropology*. Blackwell Publishing Ltd, p. 752.

Hens, S.M. & Belcastro, M.G., 2012. Auricular surface aging: a blind test of the revised method on historic Italians from Sardinia. *Forensic science international*, 214(1–3), p.209.e1-5.

Hillier, M.L. & Bell, L.S., 2007. Differentiating human bone from animal bone: a review of histological methods. *Journal of forensic sciences*, 52(2), pp.249–63.

Hillson, S., 1996. *Dental anthropology*, Cambridge University Press.

Hillson, S., 2001. Recording dental caries in archaeological human remains. *International Journal of Osteoarchaeology*, 11(4), pp.249–289.

Hillson, S., 2005. *Teeth*, Cambridge University Press.

Hillson, S. et al., 2006. The ages at death. *In: Trinkaus, E and Svoboda, J, (eds.) Early Modern Human Evolution in Central Europe. The People of Dolní Věstonice and Pavlov. (pp. 31-45). Oxford University Press: Oxford. (2006)* .

Holobinko, A., 2012. Forensic human identification in the United States and Canada: a review of the law, admissible techniques, and the legal implications of their application in forensic cases. *Forensic science international*, 222(1–3), p.394.e1-13.

Hoppa, R.D. & Vaupel, J.W., 2002. *Paleodemography* R. D. Hoppa & J. W. Vaupel, eds., Cambridge: Cambridge University Press.

Hoppitt, L. et al., 2010. Cognitive mechanisms underlying the emotional effects of bias modification. *Applied Cognitive Psychology*, 24(3), pp.312–325.

Hudson, J. & Nelson, K., 1986. Repeated encounters of a similar kind: Effects of familiarity on children's autobiographic memory. *Cognitive Development*, 1(3), pp.253–271.

Igarashi, Y. et al., 2005. New method for estimation of adult skeletal age at death from the morphology of the auricular surface of the ilium. *American Journal of Physical Anthropology*, 128(2), pp.324–339.

Inman, K. & Rudin, N., 2002. The origin of evidence. *Forensic science international*, 126(1), pp.11–6.

Innocence project, 2017. The Causes of Wrongful Conviction — The Innocence Project. Available at: http://www.innocenceproject.org/causes-wrongful-conviction [Accessed April 17, 2016].

Johansson, P. et al., 2006. How something can be said about telling more than we can know: On choice blindness and introspection.

Kassin, S.M. et al., 2010. Police-induced confessions, risk factors, and recommendations: looking ahead. *Law and human behavior*, 34(1), pp.49–52.

Kassin, S.M., 2012. Why confessions trump innocence. *The American psychologist*, 67(6), pp.431–45.

Kassin, S.M., 2014. False confessions: Causes, consequences, and implications for reform. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), pp.112-121.

Kassin, S.M., Dror, I.E. & Kukucka, J., 2013. The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, 2(1), pp.42–52.

Kassin, S.M., Dror, I.E. & Kukucka, J., 2013. The forensic confirmation bias: Problems,

perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, 2(1), pp.42–52.

Kassin, S.M. & Fong, C.T., 1999. "I'm Innocent!": Effects of Training on Judgment of Thruth and Deception in the Interrogation Room. *Law and HUman Behaviouer*, 23(5), pp.499–516.

Kassin, S.M. & Kiechel, K.L., 1996. The social psychology of false confessions: Compliance, Internalization, and Confabulation. *Psychological Science*, 7(3), pp.125–128.

Kassin, S.M. & Neumann, K., 1997. On the power of confession evidence: an experimental test of the fundamental difference hypothesis. *Law and human behavior*, 21(5), pp.469–84.

Kellman, P.J. et al., 2014. Forensic Comparison and Matching of Fingerprints: Using Quantitative Image Measures for Estimating Error Rates through Understanding and Predicting Difficulty. *Public Library of Science*, 9(5), pp.1–14.

Kemkes-Grottenthaler, A., Löbig, F. & Stock, F., 2002. Mandibular ramus flexure and gonial eversion as morphologic indicators of sex. *Homo : internationale Zeitschrift fur die vergleichende Forschung am Menschen*, 53(2), pp.97–111.

Kerkhoff, W. et al., 2015. Design and results of an exploratory double blind testing program in firearms examination. *Science & Justice*.

Kerstholt, J. et al., 2010. Does suggestive information cause a confirmation bias in bullet comparisons? *Forensic science international*, 198(1–3), pp.138–42.

Key, C.A., Aiello, L.C. & Molleson, T., 1994. Cranial suture closure and its implications for age estimation. *International Journal of Osteoarchaeology*, 4(3), pp.193–207.

Khaneman, D. & Frederick, S., 2002. Representativeness Revisited:Attribute Substitution in Intuitive Judgment. In T. Gilovich, D. Griffin, & D. Khaneman, eds. *Heruistic and biases the psychology of intuitive judgment*. Cambridge University Press, pp. 49–81.

Kimmerle, E.H., Ross, A. & Slice, D., 2008. Sexual Dimorphism in America: Geometric Morphometric Analysis of the Craniofacial Region. *Journal of Forensic Sciences*, 53(1), pp.54–57.

Klales, A.R., 2013. Current Practices in Physical Anthropology for Sex Estimation in Unidentified, Adult Individuals. In *The 82nd Annual Meeting of the American Association of Physical Anthropologists, Knoxville, Tennessee*.

Klales, A.R. & Lesciotto, K.M., 2016. The "Science of Science": Examining Bias in Forensic Anthropology. In *American Acossiation of Forensic Scince Annual Meeting*. pp. 188–189.

Klales, A.R., Ousley, S.D. & Vollner, J.M., 2012. A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. *American journal of physical anthropology*, 149(1), pp.104–14.

Klatzky, R.L., 1975. *Human memory: Structures and processes.*, W. H. Freeman.

Klayman, J., 1995. Varieties of Confirmation Bias. *Psychology of Learning and Motivation*, 32, pp.385–418.

Konigsberg, L.W., Algee-Hewitt, B.F.B. & Steadman, D.W., 2009. Estimation and evidence in forensic anthropology: Sex and race. *American Journal of Physical Anthropology*, 139(1), pp.77–90.

Krane, D.E. et al., 2008. Sequential unmasking: a means of minimizing observer effects in forensic DNA interpretation. *Journal of forensic sciences*, 53(4), pp.1006–7.

Kranioti, E. & Paine, R., 2011. Forensic anthropology in Europe: an assessment of current status and application. *Journal of anthropological sciences = Rivista di antropologia : JASS*, 89, pp.71–92.

Kranioti, E.F. et al., 2009. Sex identification and software development using digital femoral head radiographs. *Forensic Science International*, 189(1–3), p.113.e1-113.e7.

Krishan, K. et al., 2016. A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. *Forensic Science International*, 261, p.165.e1-165.e8.

Kukucka, J. & Kassin, S.M., 2014. Do confessions taint perceptions of handwriting evidence? An empirical test of the forensic confirmation bias. *Law and Human Behavior*, 38(3), pp.256–270.

Langenburg, G., 2017. Addressing potential observer effects in forensic science: a perspective from a forensic scientist who uses linear sequential unmasking techniques. *Australian Journal of Forensic Sciences*, pp.1–16.

Langenburg, G., Champod, C. & Wertheim, P., 2009. Testing for Potential Contextual Bias Effects During the Verification Stage of the ACE-V Methodology when Conducting Fingerprint Comparisons. *Journal of Forensic Sciences*, 54(3), pp.571–582.

Langley-Shirley, N. & Jantz, R.L., 2010. A Bayesian Approach to Age Estimation in Modern Americans from the Clavicle. *Journal of Forensic Sciences*, 55(3), pp.571–583.

Law Comission, The Law Commission & Law Comission, 2011. *Expert Evidence in Criminal Proceedings in England and Wales*,

Leippe, M.R., 1980. Effects of integrative memorial and cognitive processes on the correspondence of eyewitness accuracy and confidence. *Law and Human Behavior*, 4(4), pp.261–274.

Levine, L.J., 1997. Reconstructing Memory for Emotions. *Journal of Experimental Psychology: General*, 126(2), pp.165–177.

Lewis, C.J. & Garvin, H.M., 2016. Reliability of the Walker Cranial Nonmetric Method and Implications for Sex Estimation. *Journal of Forensic Sciences*, 61(3), pp.743–51.

Lord, C.G. et al., 1979. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11), pp.2098–2109.

Lottering, N. et al., 2013. Evaluation of the Suchey-Brooks method of age estimation in an Australian subpopulation using computed tomography of the pubic symphyseal surface. *American journal of physical anthropology*, 150(3), pp.386–99.

Lovejoy, C.O. et al., 1985. Chronological metamorphosis of the auricular surface of the ilium: A new method for the determination of adult skeletal age at death. *American Journal of Physical Anthropology*, 68(1), pp.15–28.

Mahakkanukrauh, P. et al., 2016. Osteometric sex estimation from the os coxa in a Thai population. *Forensic Science International*, 0(0), pp.119–127.

Mahfouz, M. et al., 2007. Patella sex determination by 3D statistical shape models and nonlinear classifiers. *Forensic Science International*, 173(2–3), pp.161–170.

Margot, P., 2011. Forensic science on trial - What is the law of the land? *Australian Journal of Forensic Sciences*, 43(2–3).

Márquez-Grant, N., 2015. An overview of age estimation in forensic anthropology: perspectives and practical considerations. *Annals of human biology*, 42(4), pp.1–15.

Martins, R., Oliveira, P.E. & Schmitt, A., 2012. Estimation of age at death from the pubic symphysis and the auricular surface of the ilium using a smoothing procedure. *Forensic science international*, 219(1–3), p.287.e1-7.

Martire, K.A. et al., 2013. The expression and interpretation of uncertain forensic science evidence: Verbal equivalence, evidence strength, and the weak evidence effect. *Law and Human behavior*, 37(3), pp.197–207.

Marx, R.B., Hochrein, M.J. & Fasano, A.D., 2014. Chapter 4 – More Pieces of the Puzzle: F.B.I. Evidence Response Team Approaches to Scenes with Commingled Evidence. In *Commingled Human Remains*. pp. 57–85.

Mattijssen, E.J.A.T. et al., 2016. Implementing context information management in forensic casework: Minimizing contextual bias in firearms examination. *Science and Justice*, 56(2), pp.113–122.

Meissner, C.A. & Kassin, S.M., 2002. "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior*, 26(5), pp.469–480.

Mennell, J. & Shaw, I., 2006. The future of forensic and crime scene science. Part I. A UK forensic science user and provider perspective. *Forensic science international*, 157 Suppl, pp.S7-12.

Merritt, C.E., 2015. The influence of body size on adult skeletal age estimation methods. *American Journal of Physical Anthropology*, 156(1), pp.35–57.

Meyer, M.A. & Booker, J.M., 1991. *Eliciting and Analyzing Expert Judgment: A Practical Guide*, SIAM.

Miller, L.S., 1984. Bias among forensic document examiners: A need for procedural change. *Journal of Police Science & Administration*, 12(4), pp.407–411.

Milner, G.R. & Boldsen, J.L., 2012. Transition analysis: a validation study with known-age modern American skeletons. *American journal of physical anthropology*, 148(1), pp.98–110.

Ministry of Justice, 2013. *The Criminal Procedure Rules Part 33 Expert Evidence*,

Mnookin, J.L., Cole, S.A., et al., 2011. The Need for a Research Culture in the Forensic Sciences. *UCLA LAW REVIEW*, 58, pp.725–779.

Morgan, R.M. & Bull, P.A., 2007. The philosophy, nature and practice of forensic sediment analysis. *Progress in Physical Geography*, 31(1), pp.43–58.

Morgan, R.M. & Bull, P.A., 2007. Foresic geoscience and crime detection, Identification, interpretation and presentation in forensic geoscience. *Minerva Mediolegale*, 127(2), pp.73–89.

Morgan, R.M. et al., 2009. The relevance of the evolution of experimental studies for the interpretation and evaluation of some trace physical evidence. *Science & justice : journal of the Forensic Science Society*, 49(4), pp.277–85.

Morgan, R.M. et al., 2014. Experimental forensic studies of the preservation of pollen in vehicle fires. *Science & justice : journal of the Forensic Science Society*, 54(2), pp.141–5.

Mulhern, D.M. & Jones, E.B., 2005. Test of revised method of age estimation from the auricular surface of the ilium. *American Journal of Physical Anthropology*, 126(1), pp.61–65.

Murrie, D.C. et al., 2013. Are forensic experts biased by the side that retained them? *Psychological science*, 24(10), pp.1889–97.

Musseweiler, T. & Englich, B., 2005. Subliminal anchoring: Judgmental consequences and underlying mechanisms. *Organizational Behavior and Human Decision Processes*, 98, pp.133–143.

Naikmasur, V.G., Shrivastava, R. & Mutalik, S., 2010. Determination of sex in South Indians and immigrant Tibetans from cephalometric analysis and discriminant functions. *Forensic Science International*, 197(1–3), p.122.e1-122.e6.

Nakhaeizadeh, S., Dror, I.E. & Morgan, R.M., 2014. Cognitive bias in forensic anthropology: visual assessment of skeletal remains is susceptible to confirmation bias. *Science & Justice*, 54(3), pp.208–14.

Nakhaeizadeh, S., Dror, I.E. & Morgan, R.M., 2015. The Emergence of Cognitive Bias in Forensic and Criminal Investigations. *British Journal of American Legal Studies*, 4(2), pp.527–554.

Nakhaeizadeh, S., Hanson, I. & Dozzi, N., 2014. The power of contextual effects in forensic anthropology: a study of biasability in the visual interpretations of trauma analysis on skeletal remains. *Journal of forensic sciences*, 59(5), pp.1177–83.

Nakhaeizadeh, S. & Morgan, R.M., 2015. Forensic Anthropology and Cognitive Bias. In A. Jamieson & M. A.A, eds. *Wiley Encyclopedia of Forensic Science*. Chichester, UK: John Wiley.

National Acadamy of Science & National Research Council, 2009. *Strengthening Forensic Science in the United States: A Path Forward*, Washington D.C.

Neisser, U., 1976. *Cognition and Reality: Principles and Implications of Cognitive Psychology*, W.H. Freeman.

Nickerson, R.S., 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), pp.175–220.

Nisbett, R.E. & Wilson, T.D., 1977. Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, (84), pp.231–259.

O'Brian, B., 2009. Prime suspect: an examination of factors that aggravate and counteract confirmation bias in criminal investigations. *Psychology, Public Policy and Law*, 15(4), pp.315–334.

Ordóñez, L. & Benson, L., 1997. Decisions under Time Pressure: How Time Constraint Affects Risky Decision Making. *Organizational Behavior and Human Decision Processes*, 71(2), pp.121–140.

Osborne, N.K.P. et al., 2014. Does contextual information bias bitemark comparisons? *Science & Justice*, 54(4), pp.267–73.

Osborne, N.K.P., Taylor, M.C. & Zajac, R., 2016. Exploring the role of contextual information in bloodstain pattern analysis: A qualitative approach. *Forensic Science International*, 260, pp.1–8.

Page, M., Taylor, J. & Blenkin, M., 2012. Context effects and observer bias--implications for forensic odontology. *Journal of forensic sciences*, 57(1), pp.108–12.

Passalacqua, N. V., 2009. Forensic Age-at-Death Estimation from the Human Sacrum. *Journal of Forensic Sciences*, 54(2), pp.255–262.

Passalacqua, N. V & Fenton, T.W., 2012. Developments in Skeletal Trauma : Blunt-Force Trauma. In D. C. Dirkmaat, ed. *A Companion to Forensic Anthropology*. Blackwell Publishing Ltd, pp. 400–411.

Patriquin, M.L., Loth, S.R. & Steyn, M., 2003. Sexually dimorphic pelvic morphology in South African whites and blacks. *HOMO - Journal of Comparative Human Biology*, 53(3), pp.255–262.

Perillo, J.T. & Kassin, S.M., 2011. Inside interrogation: The lie, the bluff, and false confessions. *Law and Human Behavior*, 35(4), pp.327–337.

Petraco, N.D.K. et al., 2012. Addressing the National Academy of Sciences' Challenge: A Method for Statistical Pattern Comparison of Striated Tool Marks. *Journal of Forensic Sciences*, 57(4), pp.900–911.

Phenice, T.W., 1969. A newly developed visual method of sexing the os pubis. *American journal of physical anthropology*, 30(2), pp.297–301.

Phillips, M.R. et al., 1999. Double-blind photography administration as a safeguard against investigator bias. *Journal of Applied Psychology*, 84(6), pp.940–951.

Pickering, R.B. & Bachman, D., 2009. *The Use of Forensic Anthropology:*, Boca Raton.

Pinheiro, J., Cunha, E. & Symes, S., 2015. Over-interpretation of bone injuries and implications for cause and manner of death. In *Skeletal trauma analysis*. Chichester, UK: John Wiley & Sons, Ltd, pp. 27–41.

Popper, K., 1959. *The Logic of Scientific Discovery*, London: Hutchinson.

Porter, S., Woodworth, M. & Birt, A.R., 2000. Truth, lies, and videotape: an investigation of the ability of federal parole officers to detect deception. *Law and human behavior*, 24(6), pp.643–58.

President's Committee of Advisors on Science and Technology, 2016. *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*,

Pretorius, E., Steyn, M. & Scholtz, Y., 2006. Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. *American Journal of Physical Anthropology*, 129(1), pp.64–70.

Redmayne, M. et al., 2011. Forensic science evidence in question. *Criminal Law Reviuew*, (5), pp.347–356.

Rhine, S., 1990. Nonmetric Skull Racing. In G. W. Gill & S. Rhine, eds. *Skeletal attribution of race*. Maxwell Museum of Anthropology, pp. 9–20.

Rholes, W.S., Riskind, J.H. & Lane, J.W., 1987. Emotional states and memory biases: effects of cognitive priming and mood. *Journal of personality and social psychology*, 52(1), pp.91–9.

Ribaux, O., Crispino, F. & Roux, C., 2015. Forensic intelligence: deregulation or return to the roots of forensic science? *Australian Journal of Forensic Sciences*, 47(1).

Ricciardi, L. & Demos, M., 2015. *Making a Murderer*, United States: Netflix.

Risinger, D.M. et al., 2002. The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion. *California Law Review*, 90(NO 1), pp.3–56.

Rissech, C. et al., 2007. Estimation of Age-at-Death for Adult Males Using the Acetabulum, Applied to Four Western European Populations. *Journal of Forensic Sciences*, 52(4), pp.774–778.

Roediger, H.L. & McDermott, K.B., 1995. Creating False Memories: Remembering Words Not Presented. *Journal of Experimental Psychology:Learning Memory and Cognition*, 21(4), pp.803–804.

Rogers, T. & Saunders, S., 1994. Accuracy of sex determination using morphological traits of the human pelvis. *Journal of forensic sciences*, 39(4), pp.1047–56.

Rubin, K.M. & DeLeon, V.B., 2017. Ancestral Variation in Orbital Rim Shape: A Three-Dimensional Pilot Study. *Journal of Forensic Sciences*.

Safer, M.A. & Keuler, D.J., 2002. Individual differences in misremembering pre-psychotherapy distress: personality and memory distortion. *Emotion (Washington, D.C.)*, 2(2), pp.162–78.

Saks, M.J. et al., 2003. Context effects in forensic science: a review and application of the science of science to crime laboratory practice in the United States. *Science & justice : journal of the Forensic Science Society*, 43(2), pp.77–90.

Scherr, K.C., Miller, J.C. and Kassin, S.M., 2014. “Midnight Confessions”: The Effect of Chronotype Asynchrony on Admissions of Wrongdoing. *Basic and Applied Social Psychology*, *36*(4), pp.321-328.

Scheuer, L., 2002. Application of osteology to forensic medicine. *Clinical Anatomy*, 15(4), pp.297–312.

Schiffer, B. & Champod, C., 2007. The potential (negative) influence of observational biases at the analysis stage of fingermark individualisation. *Forensic science international*, 167(2–3), pp.116–20.

Scott, K.R. et al., 2014. The transferability of diatoms to clothing and the methods appropriate for their collection and analysis in forensic geoscience. *Forensic science international*, 241, pp.127–37.

Silverman, B., 2011. *Research and Development in Forensic Science : a Review*, Simon, H.A., 1957. *Models of man; social and rational.*,

Smit, N.M. et al., 2016. Using Bayesian networks to guide the assessment of new evidence in an appeal case. *Crime Science*, 5(1), p.9.

Smith, M.J., 2012. A study of Interobserver Variation in Cranial Measurments and the Resulting Consequences when Analysed uring CranID. In P. D. Mitchell & J. Buckberry, eds. *Proceedings of the twelfth annual conference of the British Association for Biological Anthropology and Osteoarchaeology : Department of Archaeology and Anthropology, University of Cambridge 2010*. Archaeopress, p. 181.

Snyder, M. & Cantor, N., 1979. Testing hypotheses about other people: The use of historical knowledge. *Journal of Experimental Social Psychology*, 15(4), pp.330–342.

Soni, G., Dhall, U. & Chhabra, S., 2010. Determination of Sex From Femur: Discriminant Analysis. *Journal of Anatomical Society of India*, 59(2), pp.216–221.

Sorg, M. & Haglund, W., 2001. Advancing forensic taphonomy: purpose, theory and practice. In W. D. Haglund & M. H. Sorg, eds. *Advances in Forensic Taphonomy: Method, Theory, and Archaeological Perspectives*. CRC Press, pp. 3–29.

Soukara, S. et al., 2009. What really happens in police interviews of suspects? Tactics and confessions. *Psychology, Crime & Law*, 15(6), pp.493–506.

Sperber, D., Cara, F. & Girotto, V., 1995. Relevance theory explains the selection task. *Cognition*, 57(1), pp.31–95.

Spradley, K.M. & Jantz, R.L., 2016. Ancestry Estimation in Forensic Anthropology: Geometric Morphometric versus Standard and Nonstandard Interlandmark Distances. *Journal of Forensic Sciences*, 61(4), pp.892–897.

Spradley, M.K. et al., 2008. Demographic Change and Forensic Identification: Problems in Metric Identification of Hispanic Skeletons. *Journal of Forensic Sciences*, 53(1), pp.21–28.

Spradley, M.K. & Jantz, R.L., 2011. Sex estimation in forensic anthropology: skull versus postcranial elements. *Journal of forensic sciences*, 56(2), pp.289–96.

Springer, V.A., 2007. Expectancy Effects in Forensic Evidence Handling: Social Psychological Perspectives. *The Journal of the Institute of Justice & International Studies*, 7, pp.311–322.

Stacy, R.M., 2006. Report on the Erroneous Fingerprint Individualization in the Madrid Train Bombing Case. *Jpurnal of Forensic Identification*, 56(6), pp.706–718.

Stanovich, K.E., 2014. Assessing Cognitive Abilities: Intelligence and More. *Journal of Intelligence*, 2, pp.8–11.

Steadman, D.W., Adams, B.J. & Konigsberg, L.W., 2006. Statistical basis for positive identification in forensic anthropology. *American Journal of Physical Anthropology*, 131(1), pp.15–26.

Sternberg, R.J., 2002. *Why smart people can be so stupid*, Yale University Press.

Steyn, M., Pretorius, E. & Hutten, L., 2004. Geometric morphometric analysis of the greater sciatic notch in South Africans. *Homo : internationale Zeitschrift fur die vergleichende Forschung am Menschen*, 54(3), pp.197–206.

Stoel, R.D., Dror, I.E. & Miller, L.S., 2014a. Bias among forensic document examiners: Still a need for procedural changes. *Australian Journal of Forensic Sciences*, 46(1), pp.91–97.

Stoel, R.D., Dror, I.E. & Miller, L.S., 2014b. Bias among forensic document examiners: Still a need for procedural changes. *Australian Journal of Forensic Sciences*, 46(September 2013), pp.91–97.

Sutherland, L.D. & Suchey, J.M., 1991. Use of the ventral arc in pubic sex determination. *Journal of forensic sciences*, 36(2), pp.501–11.

SWGANTH, 2016. The Scientific Working Group for Forensic Anthropology(SWGANTH) - Home. Available at: http://www.swganth.org/index.html;

Symes, S.A. et al., 2012. Traumatic Injury to Bone in Medicolegal investigations. In D. C. Dirkmaat, ed. *A companion to forensic anthropology*. Blackwell Publishing Ltd, pp. 340–389.

Tangen, J.M., Thompson, M.B. & McCarthy, D.J., 2011. Identifying fingerprint expertise. *Psychological science*, 22(8), pp.995–7.

Thompson, W., 2009a. Interpretation: Observer Effects. In *Wiley Encyclopedia of Forensic Science*. Chichester, UK: John Wiley & Sons, Ltd.

Thompson, W., 2009b. Painting the target around the matching profile: the Texas sharpshooter fallacy in forensic DNA interpretation. *Law, Probability and Risk*, 8(3), pp.257–276.

Thompson, W.C., 2009. Beyond Bad Apples: Analyzing the Role of Forensic Science in Wrongful Convictions. *Southwestern University Law Review*, 37, pp.971–994.

Thompson, W.C., 2011. What Role Should Investigative Facts Play in the Evaluation of Scientific Evidence. *Australian Journal of Forensic Sciences*, 43(2–3), pp.123–134.

Thompson, W.C. & Cole, S.A., 2007. Psychological Aspects of Forensic Identification Evidence. In M. Costanzo, D. Krauss, & K. Pezdek, eds. *Expert psychological testimony for the courts*. New York: Lawrence Erlbaum & Association, pp. 31–68.

Thompson, W.C. & Ford, S., 1991. The meaning of a match: sources of ambiguaty in the interpretation of DNA prints. In J. Farley & J. Harrington, eds. *Forensic DNA technology*. New York: Lewis Publishers.

Tully, G., 2015. *Forensic Science Regulator Annual Report*,

Tversky, A. & Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Biases. *Science (New York, N.Y.)*, 185(4157), pp.1124–31.

Ubelaker, D.H., 2015. *The global practice of forensic science*,

Ubelaker, D.H. & Volk, C.G., 2002. A test of the phenice method for the estimation of sex. *Journal of forensic sciences*, 47(1), pp.19–24.

Urbanová, P. et al., 2014. Testing the reliability of software tools in sex and ancestry estimation in a multi-ancestral Brazilian sample. *Legal Medicine*, 16(5), pp.264–273.

Walker, P.., 1995. Problems of preservation and sexism in sexing: some lessons from historical collections for palaeodemographers. In S. Saunders & A. Herring, eds. *Grave* *reflections: portraying the past through cemetery studies.* Toronto: Canadian Scholars' Press, pp. 31–47.

Walker, P.L., 2008. Sexing skulls using discriminant function analysis of visually assessed traits. *American journal of physical anthropology*, 136(1), pp.39–50.

Walrath, D.E., Turner, P. & Bruzek, J., 2004. Reliability test of the visual assessment of cranial traits for sex determination. *American journal of physical anthropology*, 125(2), pp.132–7.

Weiss, K.M., 1972. On the systematic bias in skeletal sexing. *American journal of physical anthropology*, 37(2), pp.239–49.

Wells, G.L., Wilford, M.M. & Smalarz, L., 2013. Forensic science testing: The forensic filler-control method for controlling contextual bias, estimating error rates, and calibrating analysts' reports. *Journal of Applied Research in Memory and Cognition*, 2(1), pp.53–55.

Wescott, D.J. & Drew, J.L., 2015. Effect of obesity on the reliability of age-at-death indicators of the pelvis. *American Journal of Physical Anthropology*, 156(4), pp.595–605.

White, T.D. & Folkens, P.A., 2005. *The human bone manual*, Elsevier Academic Press.

Williams, B.A. & Rogers, T., 2006. Evaluating the accuracy and precision of cranial morphological traits for sex determination. *Journal of forensic sciences*, 51(4), pp.729–35.

Wilson, T.D. & Dunn, E.W., 2004. Self-Knowledge: Its Limits, Value, and Potential for Improvement. *Annual Review of Psychology*, 55(1), pp.493–518.

Winograd, E., Peluso, J.P. & Glover, T.A., 1998. Individual differences in susceptibility to memory illusions. *Applied Cognitive Psychology*, 12(7), pp.S5–S27.

Wood, B.P., 1999. Visual Expertise1. *Radiology*, 211(1), pp.1–3.

Wyer, R.S. & Srull, T.K., 1986. Human cognition in its social context. *Psychological Review*, 93(3), pp.322–359.

Yaşar Işcan, M., 1988. Rise of forensic anthropology. *American Journal of Physical Anthropology*, 31(S9), pp.203–229.

Appendix A Research Papers

Includes published articles by the author (with the first page of the research paper provided) and a list of articles under review, and in preparation.

Article in press

Dror, I.E., Morgan, R.M., Rando, C. & Nakhaeizadeh S., 2017. The Bias Snowball and the Bias Cascade Effects: Two Distinct Biases that May Impact Forensic Decision-Making. *Journal of Forensic Science*, (in press).

Nakhaeizadeh, S., Morgan, R.M., Rando, C. & Dror Itiel., 2017. Initial exposure to extraneous information at the crime scene and cascading bias in the subsequent evaluation of skeletal remains *Journal of Forensic Science*, (under reviuew).

Pending publications

Nakhaeizadeh, S., Rando, C. Morgan, R.M., & Dror Itiel., 2017. Cognitive and contextual influences in forensic anthropology: the implications of observer effect in biological profiling *Australian Journal of Forensic Sciences* (in preperation).

Earwaker, H., Nakhaeizadeh, S. Smit, N., & Morgan., R.M., 2017. A cultural change to enable improved decision-making in the forensic sciences: a six phased approach *Science and Justice* (in preparation).

Earwaker, H., Nakhaeizadeh, S. Smit, N., Morgan., R.M., & Harris., A., 2017. The role of psychological theories in the enhancement of the forensic science decision-making: An experimental study into the effect of priming *Forensic Science International* (in preperation).

Gamble, S., Nakhaeizadeh, S. Smit, N., Earwaker, H., Wilks, B., & Morgan., R.M., 2017. Perceptions of evidential weight surrounding forensic evidence *Science and Justice* (in preperation)

Appendix B Materials for Chapter 3

Includes the contextual information used for the skulls and the os coxa, participants answering sheet, consent form, as well as the data analysed (please refer to attached data CD labeled Appendix B Chapter 3)

THE CONTEXTUAL INFORMATION

Skull 8.849 FAO 90 (Skull 1 sex and ancestry)

This individual was excavated from the lower St Brides Churchyard Farringdon (Site code FAO 90) known for being a churchyard of low socio-economic status dated back to post-medieval period, with majority of individuals being of Caucasian decent. Osteological reports from the UCL Archive shows the metric methods applied on the long bones (found in association with the skull) to be indicative of a Male.

Skull 8.849 FAO 90 (Skull 1 sex and ancestry)

This individual was excavated from the lower St Brides Churchyard Farringdon (Site code FAO 90) known for being a churchyard of low socio-economic status dated back to post-medieval period, with majority of individuals being of Asian decent. Osteological reports from the UCL Archive shows the states metric methods applied on the long bones (found in associated with the skull) indicative of a Female.

Skull 8783 (Skull 2 sex and ancestry)

This individual was excavated from St Brides Crypt which is known to represent individuals of higher economic status with the majority of the burials being of Caucasian descent. According to grave context description a belt buckle (commonly used by Males) was found in addition to the skeletal remains.

Skull 8783 (Skull 2 sex and ancestry)

This individual was excavated from St Brides Crypt which is known to represent individuals of higher economic status with the majority of the burials being of Caucasian descent. According to grave context description a whalebones (commonly used for corset stays) was found in addition to the skeletal remains.

Skull 8.800 FAO 90 (Skull 3 sex and ancestry)

This skull was excavated from the Farringdon site dated back to post-medieval period. (Site code FAO 90). The Farringdon site burials are known for its low socio-economic status with poor church records. The osteological report taken from the UCL Archaeology Archive has stated that the morphological features associated with this individual from the femur is most in keeping with those of a Caucasian ancestry. The osteological report also indicates morphological features of the pelvic associated with this skull to have a narrow *Sciatic notch*, large *Sacro iliac joint*. and an inverted U shaped *Sub pubic arch*.

Skull 8.800 FAO 90 (Skull 3 sex and ancestry)

This skull was excavated from the Farringdon site dated back to post-medieval period. (Site code FAO 90). The Farringdon site burials are known for its low status with poor church records. The osteological report taken from the UCL Archaeology Archive has stated that the morphological features associated with this individual are most in keeping with those of an Asian ancestry. The osteological report also indicates morphological features of the pelvic associated with this skull to have a broad and shallow *Sciatic notch*, small *Sacro iliac joint*. and an inverted U shaped Sub pubic arch.

Pelvis 3724 (os coxa 1 sex and age at death)

This individual was excavated from a churchyard in Chichester, southern England dated back to post-medieval period. This individual was found buried together with juvenile skeletal remains with osteological records indicating age at death between 25-36 years of age.

Pelvis 3724 (os coxa 1 sex and age at death)

This individual was excavated from a churchyard in Chichester south of England dated back to post-medieval period. The metric measurements taken from the long bones are indicative of a male with age assessments on the sternal extremities indicates an older adult.

Pelvis 2355 (os coxa 2 sex and age at death)

This individual was excavated from a post-medieval churchyard in Chichester, southern England. The remains were found alongside two other male individuals. The osteological report includes sex estimations on the skull of this individual which demonstrates dominantly male morphological features with all sutures and bones fused indicating an old adult.

Pelvis 2355 (os coxa 2 sex and age at death)

This individual was excavated from a post-medieval churchyard in Chichester, southern England. The remains were found alongside two other female individuals. The osteological report includes sex estimations on the skull of this individual which demonstrates dominantly female morphological features with partly suture fusion and majority of post cranial bones fused, indicating a younger adult.

THE BIOLOGICAL PROFILE FORM

(showing only the sections used for the skulls and the os coxa)

Skull 1 (please estimate sex and ancestry for this skull)

Trait	Female	Male	Observation
Supraorbital ridge/torus	Less prominent	More prominent	
Eye orbit shape	Rounded	Square	
Cranial vault	Smaller, smother	Larger, rougher	
Chin	V-shaped, rounded	U-shaped, square	
Occipital condyles	Smaller	Larger	
Zygomatic process	Not expressed beyond zygomatic arch	Expressed beyond zygomatic arch	

Sex estimate: Female Male Undetermined

Confidence level percentage:_________

Feature	1F	2F?	3U	4M?	5M
Nuchal Crest					
Mastoid process					
Supra-orbital margin					
Glabella					

Sex estimate: Female Male Undetermined

Confidence level (percentage 1-100):_________

Final answer: Female Male Undetermined

Confidence level (percentage 1-100):__________

Feature	WHITE	BLACK	ASIAN
Brow-ridges	Heavy	Small	Small
Muscle marks	Rugged	Smooth	Smooth
Vault sutures	Simple	Simple	Complex
Post-bregma	Straight	Depressed	Straight
Profile (face)	Straight	Projecting	Intermediate
Shape (face)	Narrow	Narrow	Wide
Orbits	Angular	Rectangular	Rounded
Lower border-eye	Receding	Receding	Projecting
Root-nose	High,narrow	Low, rounded	Low, ridge
Bridge- nose	High	Low	Low
Spine- nose	Pronounced	Small	Small
Lower border-nose	Sharp (sill)	Guttered	Flat, sharp
Width-nose	Narrow	Wide	Medium

Ancestry estimate: White Black Asian Undetermined

Confidence level percentage: _________

Skull 2 (please estimate sex and ancestry for this skull)

Trait	Female	Male	Observation
Supraorbital ridge/torus	Less prominent	More prominent	
Eye orbit shape	Rounded	Square	
Cranial vault	Smaller, smother	Larger, rougher	
Chin	V-shaped, rounded	U-shaped, square	
Occipital condyles	Smaller	Larger	
Zygomatic process	Not expressed beyond zygomatic arch	Expressed beyond zygomatic arch	

Sex estimate: Female Male Undetermined

Confidence level percentage:_________

Feature	1F	2F?	3U	4M?	5M
Nuchal Crest					
Mastoid process					
Supra-orbital margin					
Glabella					

Sex estimate: Female Male Undetermined

Confidence level (percentage 1-100):_________

Final answer for all methods combined: Female Male Undetermined

Confidence level percentage (1-100):__________

Feature	WHITE	BLACK	ASIAN
Brow-ridges	Heavy	Small	Small
Muscle marks	Rugged	Smooth	Smooth
Vault sutures	Simple	Simple	Complex
Post-bregma	Straight	Depressed	Straight
Profile (face)	Straight	Projecting	Intermediate
Shape (face)	Narrow	Narrow	Wide
Orbits	Angular	Rectangular	Rounded
Lower border-eye	Receding	Receding	Projecting
Root-nose	High,narrow	Low, rounded	Low, ridge
Bridge- nose	High	Low	Low
Spine- nose	Pronounced	Small	Small
Lower border-nose	Sharp (sill)	Guttered	Flat, sharp
Width-nose	Narrow	Wide	Medium

Ancestry estimate: White Black Asian Undetermined

Confidence level (percentage 1-100): _________

Skull 3 (please estimate sex and ancestry for this skull)

Trait	Female	Male	Observation
Supraorbital ridge/torus	Less prominent	More prominent	
Eye orbit shape	Rounded	Square	
Cranial vault	Smaller, smother	Larger, rougher	
Chin	V-shaped, rounded	U-shaped, square	
Occipital condyles	Smaller	Larger	
Zygomatic process	Not expressed beyond zygomatic arch	Expressed beyond zygomatic arch	

Sex estimate: Female Male Undetermined

Confidence level percentage:_________

Feature	1F	2F?	3U	4M?	5M
Nuchal Crest					
Mastoid process					
Supra-orbital margin					
Glabella					

Sex estimate: Female Male Undetermined

Confidence level (percentage 1-100):_________

Final answer for all methods combined: Female Male Undetermined

Confidence level (percentage 1-100):__________

Feature	WHITE	BLACK	ASIAN
Brow-ridges	Heavy	Small	Small
Muscle marks	Rugged	Smooth	Smooth
Vault sutures	Simple	Simple	Complex
Post-bregma	Straight	Depressed	Straight
Profile (face)	Straight	Projecting	Intermediate
Shape (face)	Narrow	Narrow	Wide
Orbits	Angular	Rectangular	Rounded
Lower border-eye	Receding	Receding	Projecting
Root-nose	High,narrow	Low, rounded	Low, ridge
Bridge- nose	High	Low	Low
Spine- nose	Pronounced	Small	Small
Lower border-nose	Sharp (sill)	Guttered	Flat, sharp
Width-nose	Narrow	Wide	Medium

Ancestry estimate: White Black Asian Undetermined

Confidence level (percentage 1-100): _________

Os Coxa 1 (Please make a sex and age at death estimation)

Os coxa	Female	Male	Observation
Ventral arc	Present	Absent	
Subpubic concavity	Wide	Narrow	
Ischiopubic ramus ridge	Narrow	Broad	
Sciatic notch	Wide/broad	Narrow/small	

Sex estimate: Female Male Undetermined

Confidence level percentage: _________

Suchey Brook stage ___________

Auricular surface stage__________

Final age at death _____________

Confidence level percentage: _________

Os Coxa 2 (Please make a sex and age at death estimation)

Os coxa	Female	Male	Observation
Ventral arc	Present	Absent	
Subpubic concavity	Wide	Narrow	
Ischiopubic ramus ridge	Narrow	Broad	
Sciatic notch	Wide/broad	Narrow/small	

Sex estimate: Female Male Undetermined

Confidence level percentage: _________

Suchey Brook stage ___________

Auricular surface stage__________

Final age at death _____________

Confidence level percentage: _________

Additional comments

____

____

____

____

____

____

____

____

____

CONSENT FORM

Informed Consent Form for *Anthropological Biological Profiling* in Research Studies

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

This study has been approved by the UCL Research Ethics Committee (Project ID Number): **4672/001**

Thank you for your interest in taking part in this research. Before you agree to take part, the person organising the research must explain the project to you.

If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you to decide whether to join in.

Participant's Statement

- I have read the notes written above and the Information Sheet, and understand what the study involves.
- I understand that if I decide at any time that I no longer wish to take part in this project, I can notify the researchers involved and withdraw immediately.
- I consent to the processing of my personal information for the purposes of this research study.
- I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.
- I agree that the research project named above has been explained to me to my satisfaction and I agree to take part in this study.

Signed: Date:

Appendix C Materials for Chapter 4

Includes parts of the skeletal anthropology report, the background context for the control group ,as well as the data analysed (please refer to attached data CD labeled Appendix C Chapter)

Allegation	Date	Time	Examiner

Anthropology Skeletal Report (Mortuary exercise) Metropolitan Police, Case nr: FUZZ.12.14

Morphological Assessments

Sex Assessment

Skull

Trait	Female	Male	Observation
Supraorbital ridge/torus	Less prominent	More prominent	
Eye orbit shape	Rounded	Square	
Cranial vault	Smaller, smother	Larger, rougher	
Chin	V-shaped, rounded	U-shaped, square	
Occipital condyles	Smaller	Larger	
Zygomatic process	Not expressed beyond zygomatic arch	Expressed beyond zygomatic arch	

Sex estimate: F F? M M? Undetermined

Confidence level percentage:_________

Feature	1F	2F?	3U	4M?	5M
Nuchal Crest					
Mastoid process					
Supra-orbital margin					
Glabella					

Sex estimate: F F? M M? Undetermined

Confidence level (percentage 1-100):__________

Signature Date

Os coxa

Os coxa	Female	Male	Observation
Ventral arc	Present	Absent	
Subpubic concavity	Wide	Narrow	
Ischiopubic ramus ridge	Narrow	Broad	
Sciatic notch	Wide/broad	Narrow/small	

Sex estimate: F F? M M? Undetermined

Confidence level percentage: ________

Final sex estimation of all methods combined: F F? M M? Undetermined

Signature Date

Metric Assessments

Upper Extremity	Method Specifics	Female	Indeterminate	Male	Observation
Humeral head	Vertical diameter	<43	43-47	>47	
Radial head	Maximum diameter	≤21mm	22-23mm	≥24mm	

Sex estimate: F F? M M? Undetermined

Confidence level percentage: ________

Element	Female	Male	Observation
Clavicle (whole bone)	<138 mm	>150 mm	
Radial head	<26 mm	>29 mm	

Sex estimate: F F? M M? Undetermined

Confidence level percentage: ________

Lower Extremity	Female	Female?	Indeterminate	Male?	Male	Observation
Femoral head diameter	<42.5 mm	42.5-43.5	43.5-46.5 mm	46.5-47.5	>47.5 mm	

Sex estimate: F F? M M? Undetermined

Confidence level percentage: ________

Final sex estimation of all methods combined: F F? M M? Undetermined

Confidence level percentage: ________

Signature Date

Evidence Log Summary

Exhibit nr	Details

Signature Date

Short non-technical overall summary

Signature Date

BACKGROUND CONTEXT CONTROL GROUP 1

Background

This skeleton was found in a forensic crime scene search as part of an assessed exercise in forensic archeology. It was found buried at Old London Rd, Dorking, Surrey at the Juniper hall field centre property. The grave was only a few inches deep possibly indicating an illicit grave dug by an individual (s).

The skeleton was found blindfolded in almost full foetus position with the hands tied with silver tape. The skeleton was also found fully dressed in a red gown and a black cardigan with blue socks and a pair of black high heel shoes in size 36. Among the grave context jewellers such as a necklace and earing was found, along with foreigner money and some human hair.

Decision-making

1. In the light of the new information, would you change your previous decisions (assessments) on the skeletal remains?

Yes   No

2. If **Yes** what would your new decision (assessment) be? Please motivate.
   If **No** please elaborate you reasoning.

3. Would this information change your confidence level on your previous decision?

Yes   No

4. If **Yes** what would your new confidence level be on your previous answer?

BACKGROUND CONTEXT CONTROL GROUP 2

Background

This skeleton was found in a forensic crime scene search as part of an assessed exercise in forensic archeology. It was found buried at Old London Rd, Dorking, Surrey at the Juniper hall field centre property. The grave was only a few inches deep possibly indicating an illicit grave dug by an individual (s).

The skeleton was found blindfolded in almost full foetus position with the hands tied with silver tape. The skeleton was also found fully dressed wearing a burgundy coloured t-shirt with light grey jeans (size 34) long with a pair of dark grey converse size 37. Among the grave context jewellers such as a necklace and earing was found, along with foreigner money and some human hair.

Decision-making

1. In the light of the new information, would you change your previous decisions (assessments) on the skeletal remains?

Yes No

2. If **Yes** what would your new decision (assessment) be? Please motivate.
If **No** please elaborate you reasoning.

____
____
____
____
____
____
____
____
____
____

3. Would this information change your confidence level on your previous decision?

Yes No

4. If **Yes** what would your new confidence level be on your previous answer?

____
____

Appendix D Materials for Chapter 5

Includes letter to participants, participants answering sheet group A and B, as well as the data analysed (please refer to attached data CD labeled Appendix D Chapter 5)

LETTER TO PARTICIPANTS

Dear All,

I am writing to invite you to take part of a study that I am conducting as part of my PhD research at the Department of Security and Crime Science in sex estimation on skeletal remains.

The purpose of the research study is to investigate further on how confident people are in the application of non-metric methods on skeletal remains, with this specific study focusing on sex estimations of the skull and the os coxa.

The study requires participants with knowledge and background in the use of forensic anthropology/osteology methods to assess one skull and one os Coxa from one individual, by only applying visual methods for each skeletal element and make an interpretation of the findings. Participants will be asked to give a confidence level of their interpretations, which will be followed up with a short questioner.

The study will only take about 20 minutes and will be conducted [redacted] office on the 5th floor at the [redacted] between the periods of 21st-24th of June and 28th of June-1st of July.

A £50 pound amazon voucher will also be drawn for two participants completing the study.

If you are interested in taking part of this study please email sherry.nakhaeizadeh.12@ucl.co.uk or contact [redacted] for further information on time slots available for each day.

Thank you for your time and consideration and I am looking forward hearing from you.

With best wishes,
Sherry

FORENSIC ANTHROPOLOGY

Participants Answering Sheet (A)

Sherry Nakhaeizadeh

INSTRUCTIONS

Thank you for participating in this study. You are going to be asked to conduct some visual anthropological methods on the os coxa and the skull of one individual. Please read the consent form, and if you agree to take part of this study please fill in your answers under each section and feel free to use any reference materials that you wish. At the end of this study, you will be asked to provide some information with regards to your previous experience in forensic anthropology.

Informed Consent Form for Research in Forensic Anthropological Visual methods.

Thank you for your interest in taking part of this research.

If you have any questions arising from the information sheet or explanation already given to you please feel free to ask the research or the person present at the study.

Participant's statement

- I have read the notes written above and read the information sheet and understand what the study is about.
- I understand that if I decide at any time that I no longer wish to take part in this project, I can notify the researcher involved and withdraw immediately.
- I consent to the processing of my personal information for the purpose of this study.
- I understand that such information will be treated strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.
- I agree that the research project named above has been explained to me to my satisfaction and I agree to take part in this research.

Signed: **Date:**

Os Coxa

1. Please establish sex on this os Coxa by looking at the following traits. Please also indicate your confidence level in your final answer.

Traits	Male	Female	Undetermined	NA
The ventral arc				
The subpubic contour/concavity				
Medial aspect of the ishio-pubis ramus				

Scale	1	2	3	4	5	NA
Sciatic Notch						

Sex estimation: **Male Male? Female Female? Undetermined**

Confidence level (percentage 1-100) _____________

Skull

2. Please establish sex on this skull by looking at the following traits. Please also indicate your confidence level in your final answer.

Traits	1	2	3	4	5	NA
Nuchal crest						
Mastoid process						
Supra-orbital margin						
Supra-orbital ridge/glabella						
Mental Eminence						

Sex estimation: **Male Male? Female Female? Undetermined**

Confidence level (percentage 1-100) _____________

Final Sex assessment of this Individual

3. Please indicate your final sex estimation of the individual based on both observations on the Os Coxa and the skull. Please substantiate your answer.

Final sex estimation: **Male Male? Female Female? Undetermined**

Confidence level (percentage 1-100) ____________

Participants Questioner Form

Sex: Male Female

I am currently a: Student, (BSc) (MSc) (PhD) Professional

Please indicate your highest degree of level of education

Please state your educational background (e.g. Forensic Anthropology, Archeology, Osteology etc.)

Please state how many years of practice you had in the use of Forensic anthropological/Osteological methods on skeletal remains.

What method(s) within the establishment of a biological profile did your educational background focus more upon?

Metric methods Non-metric methods Both

Please indicate if you are more confident using:

Metric methods Non-metric methods Both

Please provide a short description on why you are more confident using one of the methods

Within non-metric methods, I am more confident when conducting sex assessment(s) on the:

Pelvis bone Skull Both

Please substantiate on your answer on why you are more confidant conducting sex estimations on the answer of your choice.

FORENSIC ANTHROPOLOGY

Participants Answering Sheet (B)

Sherry Nakhaeizadeh

INSTRUCTIONS

Thank you for participating in this study. You are going to be asked to conduct some visual anthropological methods on the os coxa and the skull of one individual. Please read the consent form, and if you agree to take part of this study please fill in your answers under each section and feel free to use any reference materials that you wish. At the end of this study, you will be asked to provide some information with regards to your previous experience in forensic anthropology.

Informed Consent Form for Research in Forensic Anthropological Visual methods.

Thank you for your interest in taking part of this research.

If you have any questions arising from the information sheet or explanation already given to you please feel free to ask the research or the person present at the study.

Participant's statement

- I have read the notes written above and read the information sheet and understand what the study is about.
- I understand that if I decide at any time that I no longer wish to take part in this project, I can notify the researcher involved and withdraw immediately.
- I consent to the processing of my personal information for the purpose of this study.
- I understand that such information will be treated strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.
- I agree that the research project named above has been explained to me to my satisfaction and I agree to take part in this research.

Signed: **Date:**

Skull

2. Please establish sex on this skull by looking at the following traits. Please also indicate your confidence level in your final answer.

Traits	1	2	3	4	5	NA
Nuchal crest						
Mastoid process						
Supra-orbital margin						
Supra-orbital ridge/glabella						
Mental Eminence						

Sex estimation: **Male** **Male?** **Female** **Female?** **Undetermined**

Confidence level (percentage 1-100) ____________

Os Coxa

1. Please establish sex on this os Coxa by looking at the following traits. Please also indicate your confidence level in your final answer.

Traits	Male	Female	Undetermined	NA
The ventral arc				
The subpubic contour/concavity				
Medial aspect of the ishio-pubis ramus				

Scale	1	2	3	4	5	NA
Sciatic Notch						

Sex estimation: **Male** **Male?** **Female** **Female?** **Undetermined**

Confidence level (percentage 1-100) ____________

Final Sex assessment of this Individual

3. Please indicate your final sex estimation of the individual based on both observations on the Os Coxa and the skull. Please substantiate your answer.

____
____
____
____
____
____
____
____
____

Final sex estimation: **Male Male? Female Female? Undetermined**

Confidence level (percentage 1-100) ____________

Participants Questioner Form

Sex:	Male	Female
I am currently a:	Student, (BSc) (MSc) (PhD)	Professional

Please indicate your highest degree of level of education

____

Please state your educational background (e.g. Forensic Anthropology, Archeology, Osteology etc.)

____

Please state how many years of practice you had in the use of Forensic anthropological/Osteological methods on skeletal remains.

____

What method(s) within the establishment of a biological profile did your educational background focus more upon?

Metric methods   Non-metric methods   Both

Please indicate if you are more confident using:

Metric methods   Non-metric methods   Both

Please provide a short description on why you are more confident using one of the methods

____
____
____
____
____
____

Within non-metric methods, I am more confident when conducting sex assessment(s) on the:

Pelvis bone Skull Both

Please substantiate on your answer on why you are more confidant conducting sex estimations on the answer of your choice.