Science and Justice 54 (2014) 208–214

Contents lists available at ScienceDirect

Science and Justice

journal homepage: www.elsevier.com/locate/scijus

Cognitive bias in forensic anthropology: Visual assessment of skeletal remains is susceptible to confirmation bias

Sherry Nakhaeizadeh [a,*], Itiel E. Dror [b], Ruth M. Morgan [a,b]

[a] Department of Security and Crime Science, University College London, London, United Kingdom
[b] UCL JDI Centre for the Forensic Sciences, University College London, London, United Kingdom

ARTICLE INFO

Article history:
Received 26 August 2013
Received in revised form 26 November 2013
Accepted 27 November 2013

Keyword:
Forensic science
Forensic anthropology
Cognitive bias
Cognitive forensic
Decision-making

ABSTRACT

An experimental study was designed to examine cognitive biases within forensic anthropological non-metric methods in assessing sex, ancestry and age at death. To investigate examiner interpretation, forty-one non-novice participants were semi randomly divided into three groups. Prior to conducting the assessment of the skeletal remains, two of the groups were given different extraneous contextual information regarding the sex, ancestry and age at death of the individual. The third group acted as a control group with no extraneous contextual information. The experiment was designed to investigate if the interpretation and conclusions of the skeletal remains would differ amongst participants within the three groups, and to assess whether the examiners would confirm or disagree with the given extraneous context when establishing a biological profile. The results revealed a significant biasing effect within the three groups, demonstrating a strong confirmation bias in the assessment of sex, ancestry and age at death. In assessment of sex, 31% of the participants in the control group concluded that the skeleton remains were male. In contrast, in the group that received contextual information that the remains were male, 72% concluded that the remains were male, and in the participant group where the context was that the remains were of a female, 0% of the participants concluded that the remains were male. Comparable results showing bias were found in assessing ancestry and age at death. These data demonstrate that cognitive bias can impact forensic anthropological non-metric methods on skeletal remains and affects the interpretation and conclusions of the forensic scientists. This empirical study is a step in establishing an evidence base approach for dealing with cognitive issues in forensic anthropological assessments, so as to enhance this valuable forensic science discipline.

1. Introduction

The complexity of data analysis and interpretation in forensic cases has been emphasised as one of the main issues in forensic science [1]. Concerns about the admissibility of evidence and expert witnesses have been raised extensively in regard to validation and error rates in techniques used by forensic scientists [2]. The National Academy of Science in the US and the Forensic Regulator in the UK have highlighted the review of standards and process within disciplines undertaking forensic science and underlined the potential for subjective interpretations and bias [3].

1.1. Cognitive bias

The issues of cognitive bias and its potential effects in forensic science and investigations have been increasingly discussed and described with empirical research demonstrating the effect of cognitive bias in decision-making within numerous forensic fields [4–9]. Research has shown that decision-making can be influenced by cognitive processes and cause forensic experts to modify their judgements [10,11]. The body of literature within forensic science and biasability has over the years recognised different sources and precipitators of cognitive influences, and confirmation bias in particular, that include time pressure [12], expectations [13], pre-existing beliefs [14] and motivation [15] which have been demonstrated to affect the final evaluations of forensic experts [16].

Studies within the fingerprint domain and DNA have demonstrated that experts were more likely to be biased when they were subjected to different types of extraneous contextual information [7,9]. In many cases, experts reached different conclusions on previous decisions when provided with extraneous contextual information [10]. It is clear that bias may impact data collection, analysis, interpretation and conclusions. It is therefore imperative that each forensic discipline examines the potential effects and presence of bias, and take measures to minimise them (appropriate measures, when needed, and that they are proportionate, see [17]), including the field of forensic anthropology.

1.2. Forensic anthropology

Forensic anthropology is a branch of applied physical anthropology that includes a spectrum of methods and skills modified from a multitude

* Corresponding author at: Department of Security and Crime Science, University College London, 35 Tavistock Square, London WC1H 9EZ, United Kingdom. Tel.: +44 20 7388 3660.
E-mail address: sherry.nakhaeizadeh.12@ucl.ac.uk (S. Nakhaeizadeh).

1355-0306
http://dx.doi.org/10.1016/j.scijus.2013.11.003

of disciplines [18]. In the legal context, forensic anthropology applies its methods to questions of medico-legal concerns [19]. Forensic anthropologists are trained to provide a biological profile (osteobiography) by using methods that enable them to provide information about an individual's sex, age, ancestry and stature [20], in addition to assisting in identification and the cause of death.

Recently there has been increased attention and interest in critically assessing some of the techniques used by forensic anthropologists. For example, in the US, the Daubert standard (1993) increased the drive to re-evaluate the methods applied [21,22]. Issues such as validation, biasability and error rates have been highlighted and assessed with recent research and re-evaluation of the methods used in the discipline [23,24].

However, there is still a lack of empirical studies in forensic anthropology examining cognitive influences that might arise and affect the judgement and final evaluation of the forensic anthropologist, especially when applying visual methodologies when assessing a biological profile on skeletal remains. The non-metric method relies on human examiners making a variety of subjective judgments, which could potentially result in experts being susceptible to bias. In a pilot web-based study, Nakhaeizadeh et al. [25] used pictures of skeletons to examine if participants were susceptible to cognitive biases in trauma analysis. This study examined the existence of bias using real skeletal remains, whereby physical anthropologists had to determine sex, ancestry, and age at death. The results indicated that bias played a role in such analysis.

2. Methodology

2.1. Research design

Based on previous research conducted by Dror and Charlton [7] on biasability in the fingerprint domain and the previous pilot study by Nakhaeizadeh et al. [25] on cognitive bias in trauma assessment, an experimental study was designed and developed to examine cognitive biases within forensic anthropological non-metric methods in assessments of sex, ancestry and age at death. To examine the biasability of forensic interpretation, participants within the field of physical anthropology were asked to establish a full biological profile of a skeleton. Participants were semi randomly divided into three groups where two of the groups were given extraneous contextual information before conducting the analysis. The third group was a control group that received no contextual information prior to the analysis. The experiment was designed to investigate whether the examiners would be affected by the extraneous context, when establishing the biological profile.

2.2. Materials

One adult skeleton was selected from the skeletal collections curated by the Centre for Human Bioarchaeology (CHB) at the Museum of London. The skeletal remains used for this research were excavated from St Bride's Lower Churchyard in 1990 by the Museum of London Archaeology Service (MoLAS)—(site code: FAO90, context number: 1474) and were analysed and recorded on to the Wellcome Osteological Research Database (WORD). The remains and cemetery were dated to the Post Medieval period with a date range for the part of the cemetery excavated to 1770–1849. The selection of this individual was made through the search of the online data from the WORD data accessed through the CHB website. The record for the selected individual states that the remains were a probable female, with an age range of 36–45 years, there was no gross observable pathology and bone preservation was very good. Ancestry was not recorded, the methods and applications to ascertain ancestry are not a recorded feature of the WORD database, but if morphological features are observed for an individual they will be recorded. This probable female did not exhibit any marked or pronounced features to suggest that they were anything other than Caucasian.

The skeletal remains were of a full body, and included a complete skull, and mandible, with the majority of postcranial elements presented in a good condition. This made it possible to conduct a visual biological profile on sex, ancestry and age at death determination (see Fig. 1). The skeletal remains also had some ambiguous features, where the morphological traits of the skull and pelvis showed no clear signs of female or male characteristics. This was of particular significance in this study because cognitive biases are more prevalent in ambiguous cases.

2.3. The contextual information

The extraneous contextual information provided to two of the participant groups before conducting their analysis, included elements such as DNA results indicating gender specific information, origin of the skeleton and age at death estimation. To examine if the contextual information had an effect on the judgement of the participants and their final evaluation of the remains, the contextual information provided to the two groups of participants contradicted each other in terms of sex assessment, ancestry and age at death (see Table 1). This made it possible to compare the different groups and determine whether there was a significant difference in the evaluation of the skeletal remains as a function of the contextual information they were exposed to. The third group was a control and was provided with no contextual information regarding sex, ancestry and age at death.

The contextual information also stated that the research was a collaboration between University College London, Museum of London and Law enforcement agencies. This was important so as to make it as authentic as possible for the participants, as it has been shown that if participants do not believe the contextual information provided it is not possible to assess whether there are any biasing effects in the decision-making [11]. The contextual information was audio recorded and played for each participant so as to make things as consistent as possible within each experimental condition.

2.4. Participants

All forty-one participants had experience and qualification in the field of physical anthropology, forensic pathology or osteology. Participants were not informed that the study was being undertaken to assess bias, as doing so would have impacted upon their performance. Participants were instead told that the study was to conduct a biological profile on skeletal remains from one complete individual by applying non-metric methods to analyse self-assessments and confidence level in using some of the most common techniques applied in forensic anthropology. This provided no further risk to participants, and followed standard ethical considerations and approval for incomplete disclosure of research objectives. All participants provided informed consent, and were anonymised following standard data protection protocols. Participants conducted the study over a three-month period at the Museum of London's Clore Learning Centre (CLC) provided by the Centre of Higher Education programme. Each participant was semi-randomly divided into one of the groups in order to ensure that each group had equally divided participants within each level of education, gender and professional background.

2.5. Procedure

The skeletal remains were laid out in anatomical order. Participants conducted the analysis alone with no one else present. The experiment took about 30 min to complete. However, to avoid time pressure, each participant was given up to 1 h to conduct their analysis.

After listening to the audio-recorded information, participants filled in a questionnaire about their own gender, level of education, professional bodies and general confidence level in assessing non-metric methods on skeletal remains. Participants were given access to visual methodology aid sheets for sex, ancestry and age at death estimations, a list combined from the methods used by the Museum of London including the majority of all non-metric assessments available for each stage.

Participants were asked to follow the biological profile form by following the order given, starting with assessing the sex of the individual,

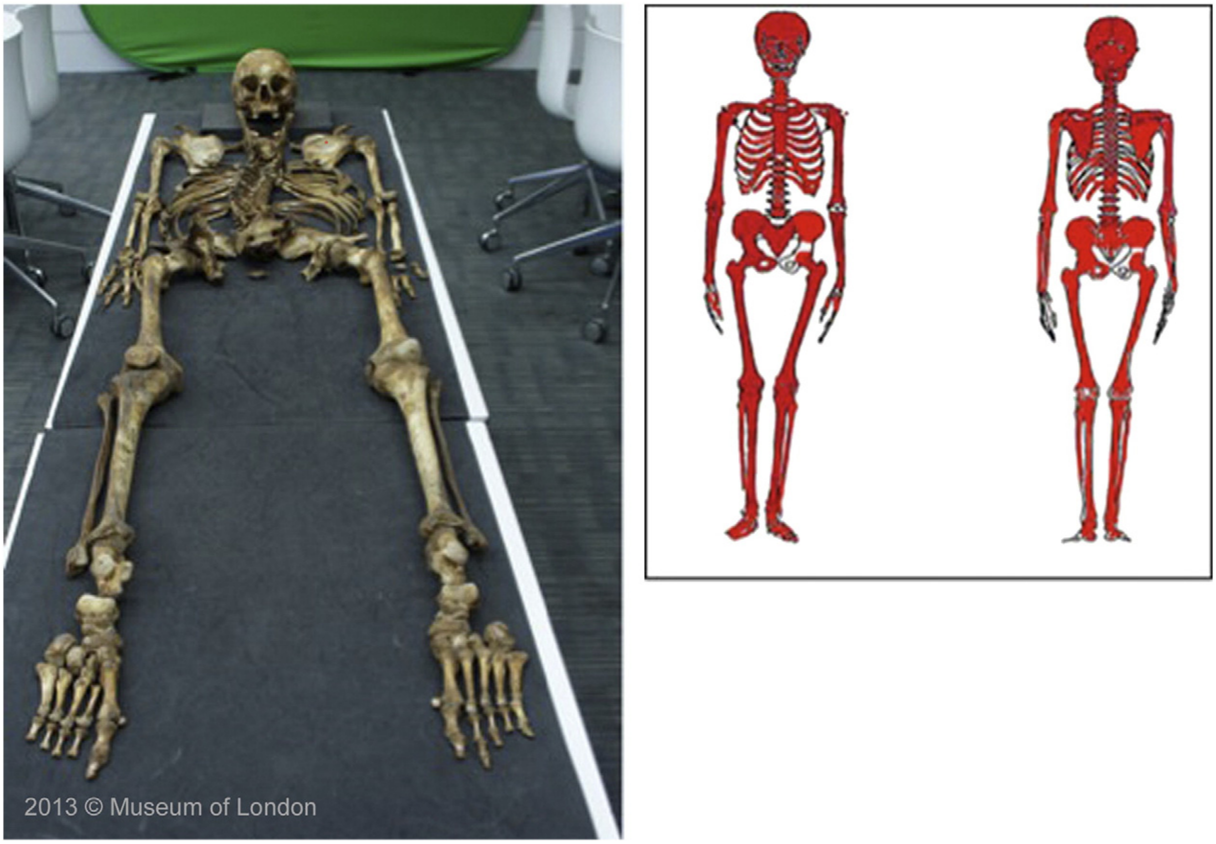

Fig. 1. Photograph of the skeletal remains used in this study (the red shading in the diagram marks the skeletal elements that are present).

followed by ancestry and age at death estimation. The age classes used in this research were adapted from the Museum of London's osteological/anthropological skeletal manual. For sex, ancestry and age at death assessments, each section was sub-divided in order to ensure that all methods possible were applied to the skeleton.

All participants were allowed to conclude "undetermined" as an option at every stage of the analysis. After each analysis conducted for sex, ancestry and age at death estimation participants were given a scale of four ranges, *Not Confident*, *Neutral*, *Confident*, and *Very Confident* to assess their level of confidence for each examination. This was included to make it more believable that the study was about confidence level and not biasability. At the end of the analysis participants were told to summarise their findings in a short sentence and place their answer charts in a sealed envelope.

2.6. Statistical analysis

The data was converted into numerical coding in order to run statistical comparison and analyse the results. The Chi-Square test was used in participants' sex, ancestry and age at death assessments to examine whether there was a significant difference between the groups as a function of the extraneous contextual information.

Table 1
Summary of the contextual information across participant groups.

Group A	Group B	Group C
Context A	Context B	Control
Gender: Male	Gender: Female	Gender: No context
Ancestry: Caucasian	Ancestry: Asian	Ancestry: No context
Age at death: 25–30	Age at death: 50–55	Age at death: No context

3. Results

3.1. Descriptive statistics

Table 2 presents the demographic and background summaries of the forty-one participants in this study.

Table 2
Summary of the participants.

	Group A	Group B	Group C
Sex			
Female	4	8	4
Male	10	6	9
Student/professional			
Student	8	9	10
Professional	6	5	3
Current/highest education			
Student MSc	4	5	9
Student PhD	4	4	1
Professional MSc	0	1	1
Professional PhD/MD	6	4	2
Occupation			
Academic/forensic anthropologist	3	3	1
Academic/forensic archaeologist	0	1	1
Academic/osteologist	1	1	1
Academic/forensic pathologist	2	0	0
Years of experience			
1–5	1	2	1
5–10	2	2	1
>10	3	1	1
General confidence level			
Not confident	0	0	0
Neutral	1	3	6
Confident	12	10	7
Very confident	1	1	0

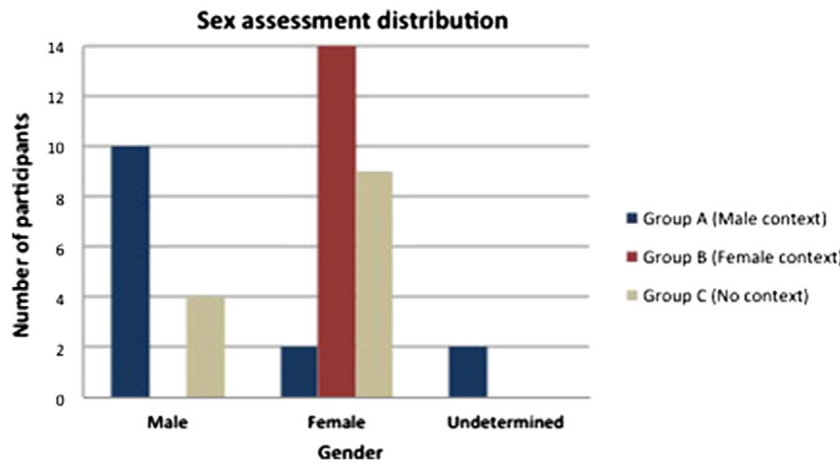

Fig. 2. Sex assessments for each group.

3.2. Comparing the groups for sex determination and biasability

There were fourteen participants in the male context (group A), fourteen in the female context (group B), and thirteen participants in the control group (group C). In the group given male context (group A), ten out of fourteen participants assessed the skeleton to be male, and four female. In the group given female context (group B), all fourteen participants assessed the skeleton to be female. In the control group (group C) four participants assessing the skeleton to be male and nine participants female (see Fig. 2).

3.2.1. Chi-Square comparing all three groups in sex assessment

The Chi-Square test was used to statistically determine whether the distribution of categorical variables between each group differed significantly from one and other. The Chi-Square test revealed a highly significant difference, with a Chi-Square, <0.000, and a p-value of <0.01 (see Table 3).

Table 3
Sex assessment analysis.

Chi-Square sex assessment	Pearson's Chi-Square	Asymp sig.
Groups A vs. B vs. C ($n = 41$)	23.037	0.000
Groups A vs. C ($n = 27$)	9.001	0.011
Groups B vs. C ($n = 27$)	5.057	0.025
Groups A vs. B ($n = 28$)	21.000	0.000

3.2.2. Chi-Square test comparing groups A (male context) and C (no context)

The result of the Chi-Square test between group A (male context) against the control group (C) with no contextual information, revealed a significant difference, with a Chi-Square, <0.011, and a p-value of <0.05, and participants sex determination, with a higher response of male estimations in group A compare to group C (see Table 3).

3.2.3. Chi-Square test comparing groups B (female context) and C (no context)

The Chi-Square test between group B (female context) against the control group (C) with no contextual information, reveals a significant difference with a Chi-Square, <0.025, and a p-value of <0.05, and participants sex determination, with higher female responses in group B compare to group C (see Table 3).

3.2.4. Chi-Square test comparing groups A (male context) and B (female context)

The Chi-Square test comparing group A (male context) against group B (female context) reveals a highly significant difference with a Chi-Square, <0.000, and a p-value of <0.01, demonstrating a higher distribution of male responses amongst participants in group A, and higher distribution of female responses in group B (see Table 3).

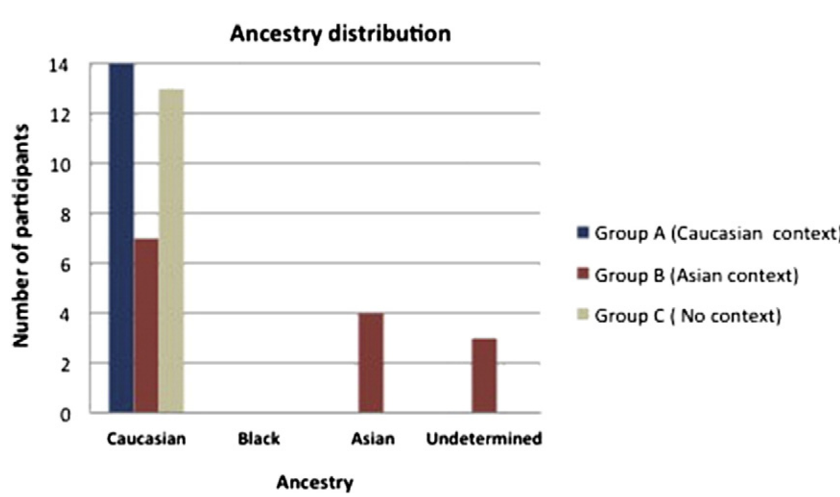

Fig. 3. Ancestry assessments for each group.

Table 4
Ancestry assessment analysis.

Chi-Square ancestry assessment	Pearson's Chi-Square	Asymp sig.
Groups A vs. B vs. C (n = 41)	16.279	0.003
Groups B vs. C (n = 27)	8.775	0.012
Groups A vs. B (n = 28)	9.333	0.009

3.3. Comparing the groups for ancestry assessment and biasability

All fourteen participants in the group given Caucasian context (group A) assessed the skeleton Caucasian. The group given Asian context (group B), seven participants assessed the skeleton to be Caucasian, four assessed Asian and three participants were undetermined in their analysis. All fourteen participants in the control group assessed the skeleton as Caucasian (see Fig. 3).

3.3.1. Chi-Square test comparing all three groups in ancestry

The Chi-Square test used to compare the counts of the categorical responses between all three groups revealed a highly significant difference between the three with a Chi-Square < 0.003 and a p-value of <0.01 (see Table 4). No Chi-Square test was applied for groups A and C because ancestry responses were in agreement for both groups, indicating no significant difference (see Table 4).

3.3.2. Chi-Square test comparing groups B (Asian context) and C (No context)

The Chi-Square test comparing group B (Asian context) against group C (no context) revealing a significant difference with a Chi-Square, <0.012, and a p-value of <0.05, between the two groups in ancestry assessments, demonstrating a higher response of *Asian* and *Undetermined* distribution amongst participants in group B, compared to group C where the responses from all participants were *Caucasian* (see Table 4).

3.3.3. Chi-Square test comparing groups A (Caucasian context) and B (Asian context)

The Chi-Square test revealed a significant difference with a Chi-Square, <0.009, and a p-value of <0.01, with all participants in group A assessing the ancestry as *Caucasian* whilst in group B there was a range of answers between *Asian*, *Caucasian* and *Undetermined* amongst participants (see Table 4).

Table 5
Age at death assessment analysis.

Chi-Square sex assessment	Pearson's Chi-Square	Asymp sig.
Groups A vs. B vs. C (n = 41)	16.705	0.010
Groups A vs. C (n = 27)	3.939	0.268
Groups B vs. C (n = 27)	5.978	0.113
Groups A vs. B (n = 28)	12.831	0.002

3.4. Comparing the groups for age at death estimation and biasability

The demographic and summary of participants' age at death assessment in each group is presented in Fig. 4.

3.4.1. Chi-Square test comparing all three groups in age at death estimation

The Chi-Square test used to investigate the distribution of categorical variables between all groups indicates a significant difference with a Chi-Square, <0.010 and a p-value of <0.05 (see Table 5).

3.4.2. Chi-Square test comparing groups A (25–30 years) and C (no context)

Chi-Square test revealed no significant difference in age at death estimation between group A (bias context) and group C (control) (see Table 5).

3.4.3. Chi-Square test comparing groups B (50–55 years) and C (no context)

The Chi-Square test revealed no significant difference in age at death estimation between group B (bias context) and group C (control) (see Table 5).

3.4.4. Chi-Square test comparing groups A (25–30 years) and B (50–55 years context)

The Chi-Square test between groups A and B indicates a significant difference with a Chi-Square, <0.002, and a p-value of <0.05, with the majority of participants age assessments in group A being 26–35 years of age and in group B being distributed between 26 and 45 and >46 years of age (see Table 5).

4. Discussion

As previous findings within the fingerprint domain and DNA have demonstrated, experts reached different conclusions on identical prints

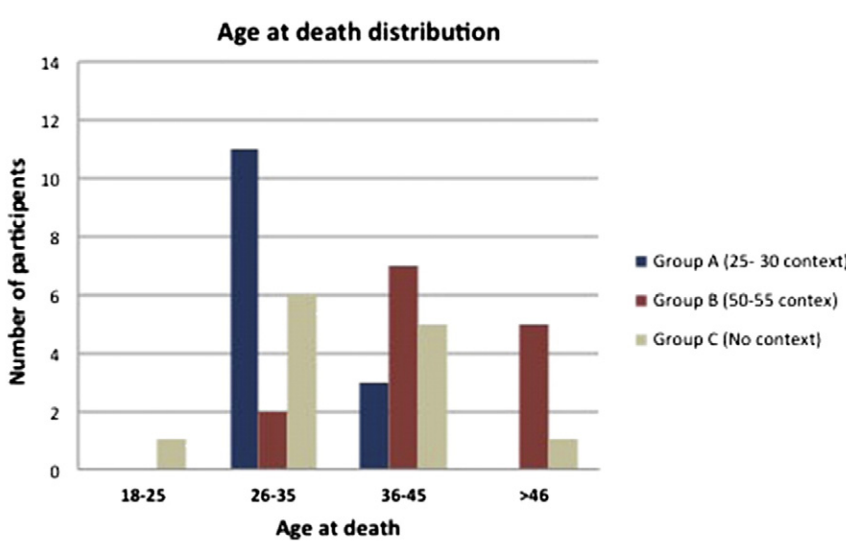

Fig. 4. Age at death assessments for each group.

and DNA samples depending on extraneous contextual information [7,9]. Similar results were found in this study. In assessment of sex, the control group was divided with 31% assessing the remains to be male and 69% female. However, in the group that received extraneous contextual information that the remains were male, 72% indicated the remains were male, 14% female and 14% undetermined in their conclusion. Of the group that was given the context that the remains were of a female, 100% of the participants concluded the remains to be female.

This clearly demonstrates that the context provided prior to anthropological assessment is of great biasing power, causing selective attention to the evidence and even overriding the actual physical evidence present. The study shows that when the data quality is ambiguous, the existence of contextual information will influence the interpretation of the participants toward an agreement with the context provided, resulting in participants focussing on characteristics of the skeleton that validate and confirm the extraneous context.

Similar results were shown in ancestry. Assessing ancestry from discrete traits is not an easy undertaking. Traditional approaches to ancestry assessments in forensic anthropology used in most forensic laboratories are based on non-empirical studies, with difficulties establishing a single known trait to be exclusively found in only one population [22]. This study demonstrates that the current methodological non metric assessments in ancestry are not only a concern in terms of their reliability, but also the issue of the methodology being vulnerable to bias interpretations. The ancestry assessments of the participants in the group given Caucasian contextual information before conducting their analysis (group A), demonstrated that 100% of the participants assessed the ancestry of the skeleton to be Caucasian. This finding was also found in the control group where 100% of the participants reached the same conclusion. However, only 50% of the participants in the group given Asian context prior to their assessment of the physical evidence (group B), assessed the ancestry to be Caucasian, the other 50% of participants were either undetermined in their decisions (21%) or, concluded the remains to be of Asian ancestry (29%).

Even though bias is most often observed in ambiguous cases, this result indicates that bias can even have an effect when the morphological characteristics are clear. The skeletal remains of the individual used in this study indicated a Caucasian ancestry of what is thought to be Caucasian features from the anthropological textbooks [22], which suggests why all participants in the control group also assessed the skeleton to be Caucasian. This shows that even though cognitive errors are less pronounced in non-ambiguous cases, it can still distort the objective processing and interpretation of data, causing some participants in the Asian given context group to arguably see what they expected to see. Comparable results were observed in Nakhaeizadeh et al. [25] studying trauma analysis where the majority of biased interpretations of trauma characteristics within the mass grave context did not only appear in the ambiguous pictures, but also in pictures illustrating clear signs of no trauma traits.

In assessment of age at death, a significant difference was observed between all three groups. There was a distinct difference between the two biasing groups (groups A and B) where there was a tendency to estimate the individual younger/older depending on the extraneous context. In group A, given contextual information that the skeletal remains were from a young adult (aged 25–30), 78% of the participants estimated the age of death to be between 26 and 35 years, in contrast to group B given contextual information regarding the individual to be 50–55 years of age, where only 14% of participants estimated 26–35 years, 50% estimated the individual to be 36–45 years and 36% estimated >46 years. In the control group (group C) 8% of the participants assessed the age at death to be 18–25 years, 46% assessed 26–35 years, 38% assessed 36–45 years and 8% assessed age at death >46 years. This not only raises the issue regarding the tendency to overestimate young individuals and underestimate older individuals when assessing age at death [19], but also that extraneous context will affect the conclusion.

The research within the field of decision-making has highlighted the dynamic and active nature of human information processing and how it makes experts distort incoming data, resulting in bias conclusions [7,10]. It has been demonstrated that these vulnerabilities are not limited to a specific field, with similar cognitive biasing issues being demonstrated across numerous forensic disciplines [11].

This study has begun to reveal the cognitive process involved in decision-making within the forensic anthropological domain. Experts make decisions on a daily basis regarding sources of evidence, and understanding the process of such judgments in forensic science is fundamental, however, this has arguably been disregarded in many domains including the field of forensic anthropology. By not considering the underlying cognitive mechanisms in force, it is possible that this will result in unproductive ways to countermeasure their outcome, which has been shown in many cases to have severe consequences [3].

Considering how science and law continue to interrelate and that the issue of scientific standards within the forensic disciplines has arisen, the forensic anthropology community must be committed to addressing these issues and providing analysis that is of the highest quality and reliability that takes these issues into account. At present, there are no agreed upon standards for the application of forensic anthropology and biological profiling, where the parameters regarding biological profiling vary considerably amongst practitioners and forensic anthropological institutions [21,24].

Nevertheless, the Scientific Working Group for Forensic Anthropology (SWGANTH) has developed consensus guidelines for the discipline, where various topic areas in forensic anthropology that could benefit from documented standardised guidelines have been acknowledged, and where limitations within non-metric methods have been highlighted. However, despite best efforts, there is still a lack in the guidelines and the body of literature within forensic anthropology regarding testing for, and dealing with, cognitive biases.

Even though a number of studies have been conducted to examine error rates within anthropological methods [22,23], the potential for cognitive biases has seldom been evaluated. This is one of the first experimental research studies in forensic anthropology to provide empirical data that has examined the impact of extraneous contextual information and confirmation biases in forensic anthropological visual assessments. The findings from this study have exposed cognitive flaws in conducting objective and impartial evaluations of skeletal remains and question the reliability of visual methodologies in ambiguous cases. Anthropological visual methodologies will continue to be needed in the future to deal with skeletal remains that are partly distorted and not ideal for metric assessments. Ideally multiple morphological features are used in assessing skeletonised remains, which is not always possible in forensic contexts [18]. The skeletal remains could be damaged due to bad preservation, and in some cases even burned and fragmented [20], resulting in complex and ambiguous scenarios. Forensic case scenarios become even more difficult when taking into account cognitive processes involved in complex decision-making.

As it has been debated [2,21], by the National Academy of Science in the US, the Forensic Regulator in the UK and recently by the UK Law Commission (2011), there are significant issues with expert opinions in court, and the presentation of erroneous information will not only bias judgements of those assessing the evidence in a case, but will also change the way in which evidence is presented during a trial which could have a major impact on the final verdict [11].

Only by revealing vulnerabilities of cognitive biases in methodologies can the forensic anthropology community develop more valid, transparent and reliable techniques where decisions will be more robust and admissible in a court of law. Therefore, there is a necessity for rigorous protocols and crucially research needs to be developed to guide these subjective interpretations and find solutions to mitigate these effects based on an empirical evidence base.

From a cognitive perspective, perceptions are never completely objective [7], and there will always be a potential of human observer error. However a more careful examination and understanding of the interaction between contextual information in the domain of forensic anthropology and biological profiling is required. Questions such as how does cognition relate to forensic anthropology, and how can it enhance forensic work needs to be evaluated. As this research has established empirically that forensic anthropologists are influenced by contextual information, further research should now use different manipulations to examine in greater complexity when such factors affect performance and render expert judgements compromised and equally when they do not, and only then will it be possible to establish appropriate counter measures.

5. Conclusion

This study identifies that decisions of forensic anthropologists based on visual assessments are vulnerable to extraneous contextual information. This study specifically demonstrated that the contextual information determined the conclusion of the forensic anthropologist in visual assessments. The results revealed a significant biasing effect, demonstrating confirmation bias within participants' assessments of sex, ancestry and age at death.

Studies within cognitive bias and forensic anthropology have an important role to play in the future development of the creditability of the discipline. At present, there is a lack of such studies that scientifically contribute to recognise the cognitive flaws and limitations within the methods applied in anthropological assessments.

This study has empirically started to identify the degree of cognitive bias and the problems it can cause within interpretations in forensic anthropological assessments, by testing for bias in non-metric methods. Cognitive bias can never be removed entirely but it can be acknowledged, restrained and thus alleviated. Within visual assessment, it is clear that the context is influencing the conclusion of the forensic anthropologist. By starting to acknowledge the power of psychological influences within forensic anthropology, careful design in technology, case management, analytical processes and blind testing can be important tools to minimise contextual biases.

Collaboration between forensic anthropology and cognitive science is essential in order to establish effective and practical workshops that focus upon critical analyses of methods and the observer effect. Equally it is important to establish solutions regarding cognitive issues, disseminated across all forensic domains. The study presented here has only begun to unravel the cognitive process involved in forensic anthropological analysis, where bias has been recognised to have an effect in visual assessments of ambiguous skeletal remains. It is clear that additional studies are necessary within the discipline in order to understand the power of cognitive bias, and how to mitigate its effect by acknowledging the weaknesses and thus strengthening the creditability of the methods, and the ability of forensic anthropology to provide robust and valuable intelligence and evidence in forensic investigations.

Acknowledgements

We wish to thank and recognise the staff at the Museum of London, Centre for Human Bioarchaeology, Jelena Bekvalac and Dr Rebecca Redfern for their support, guidance and expertise, and for making this project possible by providing us with the skeletal remains and equipment required for the experiment. We also wish to acknowledge Laura Lannin and the Centre for Higher Education for their help in providing facilities at the Museum of London's Clore Learning Centre. In addition, we would like to recognise Ally Davies for all her work and support. Our final gratitude is extended to all the participants that took their personal time to take part in the study.

References

[1] R.M. Morgan, P.A. Bull, Forensic geoscience and crime detection. Identification, interpretation and presentation in forensic geoscience, Minerva Medicoleg. 127 (2007) 73–90.
[2] J. Hiss, M. Freund, T. Kahnan, The forensic expert witness—an issue of competency, Forensic Sci. Int. 68 (2007) 89–94.
[3] I.E. Dror, S. Cole, The vision in "blind" justice: expert perception, judgment and visual cognition in forensic pattern recognition, Psychon. Bull. Rev. 17 (2010) 161–167.
[4] B. Found, F. Ganas, The management of domain irrelevant context information in forensic handwriting examination casework, Sci. Justice 53 (2013) 154–158.
[5] P. Bieber, Measuring the impact of cognitive bias in fire investigation. International symposium on fire investigation, Sci. Technol. (2012) 3–15.
[6] M. Page, J. Taylor, M. Blenkin, Context effects and observer bias—implications for forensic odontology, J. Forensic Sci. 57 (2012) 108–112.
[7] I.E. Dror, D. Charlton, Why experts make errors, J. Forensic Identif. 56 (2006) 600–616.
[8] J. Kerstholt, A. Eikelboom, T. Dijkman, R. Stoel, R.B. Van Leuven, Does suggestive information cause a confirmation bias in bullet comparisons? Forensic Sci. Int. 198 (2010) 138–142.
[9] I.E. Dror, G. Hampikian, Subjectivity and bias in forensic DNA mixture interpretation, Sci. Justice 51 (2011) 204–208.
[10] I.E. Dror, R. Rosenthal, Meta-analytically quantifying the reliability and biasability of forensic experts, J. Forensic Sci. 53 (2008) 900–903.
[11] S. Kassin, I.E. Dror, J. Kukucka, The forensic confirmation bias: problems, perspectives and proposed solutions, J. Appl. Res. Mem. Cogn. 2 (2013) 42–52.
[12] I. Evett, Evaluation and professionalism, Sci. Justice 4 (2009) 159–228.
[13] P. Bressan, M. Dal Martello, Talis pater, talis filius: perceived resemblance and the belief in genetic relatedness, Psychol. Sci. 13 (2002) 213–218.
[14] D. Hamilton, M.P. Zanna, Context effect in impression formation: changes in connotative meaning, J. Personal. Soc. Psychol. 29 (1974) 649–654.
[15] Z. Kunda, The case of motivated reasoning, Psychol. Bull. 108 (1990) 480–489.
[16] K. Ask, P.A. Granhag, Motivational sources of confirmation bias in criminal investigations. The need for cognitive closure, J. Investig. Psychol. Offender Profiling 2 (2005) 43–63.
[17] I.E. Dror, Combating bias: the next step in fighting cognitive and psychological contamination, J. Forensic Sci. 57 (2012) 276–277.
[18] S.N. Byers, Introduction to Forensic Anthropology, Third ed. Pearson Education, U.S., 2005
[19] R.B. Pickering, D.C. Bachman, The Use of Forensic Anthropology, CRC Press, Boca Raton, 2009.
[20] M. Cox, S. Mays, Human Osteology in Archaeology and Forensic Science, Greenwitch, London, 2000.
[21] A.M. Christensen, The impact of Daubert: implications for testimony and research in forensic anthropology (and the use of frontal sinuses in personal identification), J. Forensic Sci. 49 (2004) 427–430.
[22] J.T. Hefner, S.D. Ousley, D.C. Dirkmaat, Morphoscopic traits and assessment of ancestry, in: D.C. Dirkmaat (Ed.), A Comparison to Forensic Anthropology, Blackwell Publishing Ltd, 2012, pp. 287–310.
[23] C.R. Grivas, D.A. Komar, Khumo Daubert, and the nature of scientific inquiry: implications of forensic anthropology, J. Forensic Sci. 53 (2008) 771–776.
[24] M.K. Spradley, R.L. Jantz, Sex estimation in forensic anthropology: skull versus postcranial elements, J. Forensic Sci. 56 (2011) 289–296.
[25] S. Nakhaeizadeh, I. Hanson, N. Dozzi, The power of contextual effects in forensic anthropology: a study of biasability in the visual interpretations of trauma analysis on skeletal remains, J. Forensic Sci. (2014)(in press).