J Forensic Sci, 2017
doi: 10.1111/1556-4029.13569
Available online at: onlinelibrary.wiley.com

PAPER

ANTHROPOLOGY

Sherry Nakhaeizadeh,[1,2] *M.Res.; Ruth M. Morgan*,[1,2] *D.Phil.; Carolyn Rando*,[3] *Ph.D.; and Itiel E. Dror*,[1,2,4] *Ph.D.*

Cascading Bias of Initial Exposure to Information at the Crime Scene to the Subsequent Evaluation of Skeletal Remains*,†

ABSTRACT: Thirty-eight participants took part in a study that investigated the potential cascading effects of initial exposure to extraneous context upon subsequent decision-making. Participants investigated a mock crime scene, which included the excavation of clandestine burials that had a male skeletal cast dressed either in female or gender neutral clothing. This was followed by a forensic anthropological assessment of the skeletal remains, with a control group assessing the same male skeletal cast without any clothing context. The results indicated that the sex assessment was highly dependent upon the context in which participants were exposed to prior to the analysis. This was especially noticeable in the female clothing context where only one participant determined the male skeletal cast to be male. The results demonstrate the importance of understanding the role of context in forensic anthropology at an early stage of an investigation and its potential cascading effect on subsequent assessments.

KEYWORDS: forensic science, cognitive forensics, decision-making, forensic anthropology, contextual information, crime scene, bias, human factors

The central role of human cognition in forensic science and its effect on the interpretation of forensic evidence is being increasingly recognized within the forensic disciplines (1). It is clear that the concerns raised over expert decision-making and their vulnerabilities have not only been highlighted in recent key governmental reports (2–5), but also created a debate within the literature, moving toward the emergence of "cognitive forensics" as a defined field of research (6,7). Whilst the forensic community is progressively accepting the issues involved within human cognition and decision-making, the debate around how to identify, control, and minimize these effects is still ongoing (8). The issue of how to increase objectivity and to minimize cognitive biases entering a criminal investigation at an early stage has been intensified (9,10), 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 (11–15). 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 minimize task irrelevant contexts (15). Many of the recommended solutions are targeting different disciplines within forensic science (16–18) at different stages in the forensic science process (10,19,20).

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 (21). Others suggest that such exposure is good for motivating forensic examiners and for their "personal satisfaction" (22). Further concerns have been raised with regard to the fact that research into subjective decision-making might detract from focusing on increasing the objectivity with which forensic evidence can be interpreted (21), for example, through an improved understanding of the dynamics of forensic trace materials (23). Nevertheless, there are many crucial decisions being made throughout the progression of evidence from crime scene to court (24). 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, the cognitive impacts in play during the assessment of human remains have only recently begun to be assessed in published literature, with much focus being on

[1]Department of Security and Crime Science, University College London, 35 Tavistock Square, London, WC1H 9EZ, UK.
[2]Centre for the Forensic Sciences, University College London, 35 Tavistock Square, London, WC1H 9EZ, UK.
[3]Institute of Archaeology, University College London, 31-34 Gordon Square, London, WC1H 0P, UK.
[4]Cognitive Consultants International (CCI-HQ), London, UK.

*Presented at the ANZFSS 23rd International Symposium on the Forensic Sciences, September 18–22, 2016, in Auckland, New Zealand.

†Funded by the Engineering and Physical Sciences Research Council of the UK through the Security Science Doctoral Research Training Centre (UCL SECReT) based at University College London (EP/G037264/1).

Received 20 Feb. 2017; and in revised form 27 April 2017; accepted 3 May 2017.

issues of confirmation bias, expectation bias, and contextual influences (25–28). Despite the fact that contextual and environmental effects have been shown to be powerful influences on how people construct, seek, as well as interpret information (such as Anderson 2000 (29)), very little is known within forensic anthropology 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 (30), be part of the revision of search strategies (31), as well as helping to preserve, excavate, and document the skeleton *in situ* (30). 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 (23). 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, especially when making assessments on ambiguous skeletons (27).

Furthermore, studies have also identified issues for potential expectation bias when estimating sex of the skeletal remains (32,33). 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 (34,35). Despite the discussion about the effect of context in forensic anthropology, further empirical studies are required 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.

This study reports the findings from a series of experiments undertaken to investigate 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 participants in order to examine whether early exposure to such contexts would cascade (e.g., from the initial evidence collection to the evaluation and interpretation) and affect the subsequent assessment of the skeletal remains (36). This was done to further examine the extent to which contextual biases are present within forensic anthropological methods.

Methodology

Research Design

Participants in this study investigated a mock crime scene, focusing on forensic archeological 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. The research provided no risk to participants and followed standard ethical considerations and approval for incomplete disclosure of research objectives.

Materials and Context

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 arguably 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, that is, perceived as either male or female (see Fig. 1 for an example of skeletal casts dressed in female clothing (a & b) and in gender neutral garments (c & d)). This was to see whether "extraneous" clothing associated with skeletal excavations (e.g., 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 (37)). 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.

Participants

A total number of 38 MSc students participated in this 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), all with a bachelors degree and background in bioarcheology/biological and physical anthropology or osteology, with training and experience in the use of osteological techniques on skeletal remains. To minimize 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 archeology 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 archeology

FIG. 1—*Graves a and b are skeletal male casts with female clothing, and grave c and d are skeletal male casts with gender neutral clothing.*

course, with previous background knowledge of forensic anthropological/osteological assessments, were further asked to take part in the subsequent forensic anthropological analysis postexcavation. 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.

Procedures

Phase One: Preparation of the Burials and the Mock Crime Scene—Four clandestine burials were created with each grave having an approximate diameter of 120 × 80 cm, with a depth of roughly 20 cm. 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. This procedure was replicated for the following year.

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. 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. 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 archeology course. The skeletal remains were recovered and put in body bags and transported back to the mock mortuary. 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 nonmetric 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.

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 nontechnical summary of their analysis on the skeletal remains. Additionally, to understand the decision-making process further, participants were also asked to provide a confidence level (using a percentage scale 1-100) 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 (see for example Buikstra and Uberlaker (38)), as well as calipers and measurement boards for metric analysis. The time frame to conduct the analysis in the mortuary was approximately 45 min–1 h.

Phase Three: A Control Study—A control group was created, and phase three of the study was run several months after the forensic archeology module ended, with 15 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 laboratory 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 archeology 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 nontechnical 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, all 15 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 regard 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. A total of seven participants from the control group received the summary report of the female burial contexts and eight participants received the summary report of the "gender neutral clothing." 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. This was to assess whether the initial judgment of a participant would be affected by the additional information, or whether participants would confirm their initial analysis without being influenced by the additional information.

Analysis

To examine whether there was a difference between the groups as a function of the extraneous context, a series of chi-square tests were carried out at the 95% significance level (with the expected values being based on the observed values) using SPSS. In addition, due to the small sample size of some cells, a Fisher's exact test was also reported. A one-way ANOVA was conducted to compare the confidence level of the participants with the three levels of condition (female clothing, neutral clothing, and control). This was followed by Independent *t*-test to make a post hoc comparison. In addition, a dependent *t*-test was also used to compare the confidence level of participants within the control group.

Results

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 2 shows the distribution of the decisions on the skeletal remains for all three groups sex assessment with

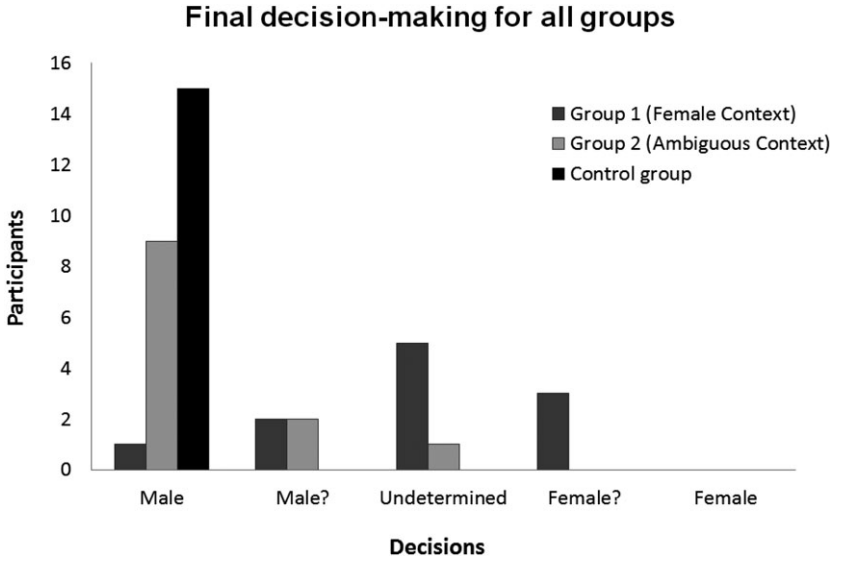

FIG. 2—*The distribution of participant's final assessment of the skeletal remains for all three groups.*

"male?" and "female?" being the representative terms that indicate "possibly" male and "possibly" female in anthropology.

Chi-Square and Fisher Exact Test Comparing Groups in Sex Assessment

Control versus Group 1 and Group 2—The chi-square 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 and Group 1 and Group 2 with a chi-square, <0.005 and a p value of <0.01. The Fisher exact test showed <0.003 and a p value of <0.01 (see Table 1 for further details).

Control versus Group 1—The chi-square 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 <0.01 and the Fisher exact test showing <0.000 and a p value of <0.01. (see Table 1 for further details).

Control versus Group 2—The chi-square was used to statistically determine whether the distribution of categorical variables between the Control group (no context) and Group 2 (gender neutral context) differed significantly from one and other. The chi-square revealed no significant difference with a chi-square, 0.121 and a p value of >0.05. This was also the case of the Fisher Exact test, showing >0.075 with a p value of >0.05 (see Table 1 for further details).

Group 1 versus Group 2—The chi-square 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 test revealed a significant difference with a chi-square, 0.007 and a p value <0.01. This was also the case of the Fisher exact test showing <0.003 and a p value of <0.01 (see Table 1 for further details).

TABLE 1—*Chi-square and Fisher's exact test results.*

	Value	df	Asymp Sig.	Exact Sig.
Control versus Group 1 and Group 2 (N = 38)				
Pearson's chi-square	12.887	3	0.005	0.002
Fisher's exact test	11.848			0.003
Control versus Group 1 (N = 26)				
Pearson's Chi-Square	22.159	3	0	0
Fisher's exact test	21.976			0
Control versus Group 2 (N = 27)				
Pearson's chi-square	4.219	2	0.121	0.075
Fisher's exact test	3.721			0.075
Group 1 versus Group 2 (N = 23)				
Pearson's chi-square	12.046	3	0.007	0.002
Fisher's exact test	11.669			0.003

TABLE 2—*An overall summary of the mean confidence value across all three groups.*

Group	N	Mean %	SD	SE
Group 1 (Female context)	9	57.78	15.635	5.212
Group 2 (Gender neutral context)	11	80.91	12.210	3.682
Control group	15	79.33	7.037	1.817
Control group after context	15	68.67	9.904	2.567

Comparing Confidence Level

A one-way repeated measure ANOVA was conducted to compare the confidence level of the participants between the three levels of conditions (female context, neutral context, and control). This was followed up by an independent 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 three groups outlined in Tables 1 and 2.

There was a significant difference in the confidence level of the participants at the $p < .05$ level for the three conditions; $F(2.32) = 12.821$, $p = 0.000$.

Crime Scene Group 1 and Group 2

The result from the independent samples t-test comparing the confidence level for participants in Group 1 (female grave context) and Group 2 (gender neutral context), indicated a statistically significant difference in the confidence level for Group 1 (M = 57.78 SD = 15.63) given female grave context and Group 2 (M = 80.91 SD = 12.21) given gender neutral grave context; $t(18) = -3.719$, $p = 0.002$.

Group 1 versus Control

The result from the independent samples t-test comparing the confidence level for participants in Group 1 (female grave context) and the Control group (no context) showed a significant difference in the confidence level for Group 1 (M = 57.78 SD = 15.63) given female grave context and the Control group (M = 79.33 SD = 7) given no context; $t(22) = -4.659$, $p = 0.000$.

Group 2 versus Control

The result from the independent samples t-test comparing the confidence level for participants in Group 2 (gender neutral context) and Control group (no context) indicated no significant difference in the confidence level for Group 2 (M = 80.91 SD = 12.21) given gender neutral grave context and the Control group (M = 79.33 SD = 7) given no grave context; $t(24)$=0.416, $p = 0.681$.

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, $p = 0.000$.

Decision-making Change of Participants in Control Group after Context

After all 15 participants assessed the male skeletal cast blindly, seven participants received the summary report of the

female burial contexts and eight participants received the summary report of the "gender neutral" burial contexts. Only two participants in total (one participant from each subgroup) 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?

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), seven 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. 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 that influence can vary and the results shows that the strength of the context in which the decision is being made, as well as the direction of the bias (39) played a significant role.

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 (40) and that once a hypothesis is 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 (41). 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 Nonmetric Assessments

This study included both metric and nonmetric analysis on nonambiguous skeletal casts, and it is important to highlight that the aim of this study was not to conduct a validation and classification study of nonmetric 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 (42,43). Furthermore, the assessment of the participants was based on the basis of the overall inferences made from all methods available, (both metric and nonmetric), 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 nonmetric 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 (44). 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 (25). Moreover, some morphological traits of skeletal casts may not be 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.

Equally, 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 nonworking 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 (45–47). 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 (10). Therefore, a valuable comparable study would be to see whether similar effects could be found amongst working professional anthropologists. In addition, observing if the same effect would have been present when dressing female casts in male clothing would be an interesting thing to study in the future.

Combining Individual Assessments

There is a growing consensus in the forensic science community with respect to cognitive biasing effects, and how to minimize the risk they pose in forensic casework. The most common solution proposed is the creation of a context free environment, separating the analyst from potentially biasing information (8), as the opinions of forensic scientists are considered to be produced independently, and not influenced by elements of the case that have no relevance to the scientific process. However, in forensic anthropology, it is important that the forensic anthropologists are present on site helping to 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 recognized that it is important to utilize 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 when developing approaches for scene management and 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. This study shows that clothing was influential and 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.

Almost every forensic case involves a variety of different specialized personnel, with both scientists and law enforcement working closely together and bringing different skills. Addressing and removing "irrelevant context" has therefore raised concerns that such an approach may create silos of the different personnel that hampers an integrated approach within the practice of forensic science and in the delivery of robust forensic reconstructions (21). Therefore, in many cases, it is acknowledged that contextual information will have a role in assisting the construction of forensic reconstructions. However, whilst not all contexts will have a biasing influence, this study does illustrate that it is possible for context to affect decision-making at a subconscious level. It is important to be aware of such instances and to ensure that inferences presented incorporate an appreciation of the potential for cascaded bias from the introduction 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 to 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 high volume laboratories within other disciplines (17,46).

As 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, and due to the complexities of the decision-making involved, which must often be made in line with existing policies or procedures, acknowledging the existence of cognitive and contextual effects and identifying situations in which it may occur is the first step in the right direction.

Moreover, improving our understanding of human decision-making within forensic anthropology (and the wider forensic science disciplines) by undertaking more empirical research, will produce data that will aid further understanding of which factors *lead* to and *influence* a decision. Embracing a constructive discussion about the role of human decision-making in the forensic sciences, and fostering a transparent and sustainable culture of context management based upon empirical findings will allow the forensic anthropology community (as well as other forensic disciplines) to openly explore decision-making within the forensic process, defining where issues exist, and finding ways in which decision-making processes can be enhanced to ensure the delivery of robust transparent forensic reconstruction approaches.

Conclusion

This study has provided an important step toward 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 only 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. The forensic anthropological community has come far in the development of the discipline, but as other disciplines have started to act and entered the dialogue of cognitive interpretation issues, further research within the discipline is needed to articulate and develop frameworks that incorporate an understanding of when context may influence the interpretation of evidence.

Acknowledgment

The authors wish to extend our thanks to Dr. Sally Gamble, Nathalie Dozzi, and Agathe Ribèreau Gayon, for their help and assistance in setting up the mock crime scene, as well as extend our thanks to all the participants who took part in this study.

References

1. Nakhaeizadeh S, Dror IE, Morgan RM. The emergence of cognitive bias in forensic and criminal investigations. Br J Am Leg Stud 2015;4 (2):527–54.
2. Government Chief Scientific Advisor. Forensic science and beyond: authenticity, provenance and assurance, evidence and case studies, 2015; https://www.gov.uk/government/publications/forensic-science-and-beyond.
3. Forensic Science Regulator. Guidance: cognitive bias effects relevant to forensic science examinations. FSR-G-217, 2015; https://www.gov.

uk/government/uploads/system/uploads/attachment_data/file/510147/217_FSR-G-217_Cognitive_bias_appendix.pdf.
4. National Commission on Forensic Science. Ensuring that forensic analysis is based upon task-relevant information, 2015; https://www.justice.gov/ncfs/file/818196/download.
5. President's Committee of Advisors on Science and Technology. Forensic science in criminal courts: ensuring scientific validity of feature comparison methods, 2016; https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf.
6. Dror IE, Stoel RD. Cognitive forensics: human cognition, contextual information, and bias. In: Bruinsma, Gerben, Weisburd, David, Editors. Encyclopedia of criminology and criminal justice. New York, NY: Springer, 2014;353–363.
7. Found B. Deciphering the human condition: the rise of cognitive forensics. Aust J Forensic Sci 2014;47(4):386–401.
8. Dror IE. Practical solutions to cognitive and human factor challenges in forensic science. Forensic Sci Policy Manag An Int J 2014;4(3–4):105–13.
9. de Gruijter M, de Poot C, Elffers H. Reconstructing with trace information: does rapid identification information lead to better crime reconstructions? J Investig Psychol Offender Profiling 2016;14(1):88–103.
10. van den Eeden CAJ, de Poot CJ, van Koppen PJ. Forensic expectations: investigating a crime scene with prior information. Sci Justice 2016;56(6):475–81.
11. Innocence Project. The causes of wrongful conviction – the innocence project 2016; http://www.innocenceproject.org/causes-wrongful-conviction.
12. Federal Bureau of Investigation. FBI — FBI testimony on microscopic hair analysis contained errors in at least 90 percent of cases in ongoing review 2015; 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.
13. Thompson WC. Beyond bad apples: analyzing the role of forensic science in wrongful convictions. Southwest Univ Law Rev 2010;1:971–94.
14. Ricciardi L, Demos M. Making a murderer. Los Gatos, CA: Netflix, 2015.
15. Dror IE, Thompson WC, Meissner CA, Kornfield I, Krane D, Saks M, et al. Letter to the Editor – Context management toolbox: a linear sequential unmasking (LSU) approach for minimizing cognitive bias in forensic decision making. J Forensic Sci 2015;60(4):1111–2.
16. Kerkhoff W, Stoel RD, Berger CEH, Mattijssen EJAT, Hermsen R, Smits N, et al. Design and results of an exploratory double blind testing program in firearms examination. Sci Justice 2015;55(6):514–9.
17. Krane DE, Ford S, Gilder JR, Inman K, Jamieson A, Koppl R, et al. Sequential unmasking: a means of minimizing observer effects in forensic DNA interpretation. J Forensic Sci 2008;53(4):1006–7.
18. Archer MS, Wallman JF. Context effects in forensic entomology and use of sequential unmasking in casework. J Forensic Sci 2016;61(5):1270–7.
19. Thompson WC. What role should investigative facts play in the evaluation of scientific evidence. Aust J Forensic Sci 2011;43(2–3):123–34.
20. Edmond G, Tangen JM, Searston RA, Dror IE. Contextual bias and cross-contamination in the forensic sciences: the corrosive implications for investigations, plea bargains, trials and appeals. Law, Probab Risk 2015;14(1):1–25.
21. Champod C. Research focused mainly on bias will paralyse forensic science. Sci Justice 2014;54(2):107–9.
22. Butt L. The forensic confirmation bias: problems, perspectives, and proposed solutions – commentary by a forensic examiner. J Appl Res Mem Cogn 2013;2(1):59–60.
23. Morgan RM, Flynn J, Sena V, Bull PA. Experimental forensic studies of the preservation of pollen in vehicle fires. Sci Justice 2014;54(2):141–5.
24. Dror IE. Cognitive neuroscience in forensic science: understanding and utilizing the human element. Philos Trans R Soc Lond B Biol Sci 2015;370(1674):20140255.
25. Klales AR, Lesciotto KM. The "science of science": examining bias in forensic anthropology. Proceedings of the 68th Annual Scientific Meeting of the American Academy of Forensic Sciences; 2016 Feb 22–27; Las Vegas, NV. Colorado Springs, CO: American Academy of Forensic Sciences, 2016.
26. Nakhaeizadeh S, Dror IE, Morgan RM. Cognitive bias in forensic anthropology: visual assessment of skeletal remains is susceptible to confirmation bias. Sci Justice 2014;54(3):208–14.
27. Nakhaeizadeh S, Hanson I, Dozzi N. 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;59(5):1177–83.
28. Passalacqua NV, De La Paz J, Zejdlik K. Identification of missing Norwegian World War II soldiers, in Karelia Russia. J Forensic Sci 2015;60(4):1104–10. *J Forensic Sci* 2016;61(5):1405–7.
29. Anderson JR. Cognitive psychology and its implications, 5th edn. New York, NY: Worth Publishers and WH Freeman, 2000.
30. Cheetham P, Hanson I. Excavation and recovery in forensic archeological investigations. In: Blaue S, Ubelaker DH, editors. Handbook of forensic anthropology and archeology. Walnut Creek, CA: Left Coast Press Inc, 2009;141–52.
31. Haglund WD. Archeology and forensic death investigations. Hist Archeol 2001;35(1):26–34.
32. Effros B, Brown TA, Brown KA, Bruzek J, Delort E, Giles E, et al. Skeletal sex and gender in Merovingian mortuary archaeology. Antiquity 2000;74(285):632–9.
33. Walker P. Problems of preservation and sexism in sexing: some lessons from historical collections for palaeodemographers. In: Saunders S, Herring A, editors. Grave reflections: portraying the past through cemetery studies. Toronto, ON: Canadian Scholars' Press, 1995;31–47.
34. Pinheiro J, Cunha E, Symes S. Over-interpretation of bone injuries and implications for cause and manner of death. In: Passalacqua N.V, Rainwater C, editors. Skeletal trauma analysis. Chichester: John Wiley & Sons, Ltd, 2015;27–41.
35. Blau S. How traumatic: a review of the role of the forensic anthropologist in the examination and interpretation of skeletal trauma. Aust J Forensic Sci 2017;49(3):261–80.
36. Dror IE, Morgan RM, Rando C, Nakhaeizadeh S. The bias snowball and the bias cascade effects: two distinct biases that may impact forensic decision making. J Forensic Sci 2017;62(3):832–3.
37. Anderson CA, Kellam KL. Belief perseverance, biased assimilation, and covariation detection: the effect of hypothetical social theories and new data. Personal Soc Psychol Bull 1992;18:555–65.
38. Buikstra JE, Ubelaker DH. Standards for data collection from human skeletal remains. proceedings of a seminar at the Field Museum of Natural History. Fayetteville, AR: Arkansas Archeological Survey, 1994.
39. Dror IE. A hierarchy of expert performance. J Appl Res Mem Cogn 2016;5:121–7.
40. Burke AS. Improving prosecutorial decision making: some lessons of cognitive science. William Mary Law Rev 2005;47(5):1587–633.
41. Anderson CA, Lepper MR, Ross L. Perseverance of social theories: the role of explanation in the persistence of discredited information. J Pers Soc Psychol 1980;39(6):1037–49.
42. Klales AR, Ousley SD, Vollner JM. A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. Am J Phys Anthropol 2012;149(1):104–14.
43. Ubelaker DH, Volk CG. A test of the phenice method for the estimation of sex. J Forensic Sci 2002;47(1):19–24.
44. Klales AR. Current practices in physical anthropology for sex estimation in unidentified, adult individuals. Am J Phys Anthropl 2013;150(S56):168.
45. Edmond G, Towler A, Growns B, Ribeiro G, Found B, White D, et al. Thinking forensics: cognitive science for forensic practitioners. Sci Justice 2017;57(2):144–54.
46. Dror IE. The paradox of human expertise: why experts get it wrong. In: Kapur N, editor. The paradoxical brain. Cambridge: Cambridge University Press, 2011;177–88.
47. Found B, Ganas J. The management of domain irrelevant context information in forensic handwriting examination casework. Sci Justice 2013;53(2):154–8.

Additional information and reprint requests:
Itiel E. Dror, Ph.D.
Department of Security and Crime Science
University College London
35 Tavistock Square
London WC1H 9EZ
UK
E-mail: i.dror@ucl.ac.uk

Appendix Cross-tab table of the group responses

			Responses				
			F?	M	M?	U	Total
Group	1 Female context	Count	3	1	2	5	11
		Expected count	0.9	7.2	1.2	1.7	11.0
	2 Gender neutral context	Count	0	9	2	1	12
		Expected count	0.9	7.9	1.3	1.9	12.0
	3 Control	Count	0	15	0	0	15
		Expected count	1.2	9.9	1.6	2.4	15.0
Total		Count	3	25	4	6	38
		Expected count	3.0	25.0	4.0	6.0	38.0