## A44 Reliability and Biasability of Sectioning-Point-Based Sex Estimation

*Stephanie Hartley, BA*, Lafayette, LA 70508; Allysha P. Winburn, PhD, University of West Florida, Pensacola, FL 32514; Itiel Dror, PhD, University College London, London, England, UNITED KINGDOM*

**Learning Overview:** After attending this presentation, attendees will understand the effects of extraneous biasing information on skeletal measurements and anthropological sex estimates.

**Impact on the Forensic Science Community:** This presentation will impact the forensic science community by showcasing that, while cognitive biases do not significantly impact measurements of the humeral and femoral heads or sex estimates of those individual skeletal elements, biasing context does statistically significantly impact the overall, categorical sex estimates of skeletal assemblages.

Scientific knowledge production is not purely objective, as it often includes subjective human cognitive processes. Forensic anthropologists must be able to make observations and draw conclusions with minimal compromise by cognitive biases in order to provide reliable and accurate expert witness testimony. A recently conducted comprehensive review of the anthropological literature revealed that there have been several studies ($n$=20) demonstrating the effect of extraneous information on morphological assessments.[1] However, no research has investigated whether anthropological metric assessments are affected by biasing context. This study was designed to assess the potential of cognitive bias in metric sex estimations.

Fifty-three trained osteologists, attendees at the 2020 *American Academy of Forensic Sciences* meeting, measured a difficult-to-classify human femur, either with or without the presence of a potentially gender-biasing photograph (a staged recovery scene with men's or women's clothing) and a strongly sexually dimorphic humerus. Participants used digital sliding calipers to measure the maximum femoral head diameter and/or vertical diameter of the head of the humerus.[2] They then used a sectioning-point sex estimation method to provide a categorical sex estimate for the individual skeletal element(s) and, if relevant, the combined skeletal assemblage.[3] The participants were randomly separated into three groups: control (given the femur only; $n$=25); female bias (given the difficult-to-classify femur, a female humerus, and female-biasing extraneous context; $n$=14); and male bias (given the difficult-to-classify femur, a male humerus, and male-biasing extraneous context; $n$=14). To minimize potential influences on the decision-making process, the participants were not informed of the full nature of this experiment; the Institutional Review Board (IRB) -approved data collection protocols also included a "distractor" ancestry estimation component. It was hypothesized that, when exposed to extraneous biasing contextual information, participants would differ in: (1) their observations (humeral and femoral head measurements); (2) their provisional sex estimation conclusions for the individual skeletal elements; and (3) their final sex estimation conclusions for the overall skeletal assemblage.

The first two hypotheses were rejected. Standard Error Index values were low for all measurements, across all groups (0.31 to 1.70mm). Krippendoff's Alpha values were high for all sex estimates, across the groups, ranging from 0.747 to 0.847 (i.e., moderate to very consistent). Student's $t$-tests indicated no statistically significant differences in the measurements taken of the femoral heads, regardless of exposure to biasing context (all $p$-values >0.05). Wilcoxon rank-sum tests indicated no statistically significant differences among the control, female-biased, and male-biased groups in their provisional sex estimates of the individual elements (all $p$-values >0.05).

The third hypothesis was supported. Wilcoxon rank-sum tests indicated statistically significant differences in overall sex estimates between the female-biased group and both the control and male-biased groups ($p$<0.001). In spite of the fact that the difficult-to-classify femur actually originated from a male anatomical skeleton, 78.5% of the female-biased group's conclusions matched the biasing information given (i.e., "Female" or "Probable Female'). In contrast, 85.7% of the male-biased group's conclusions matched the biasing information given (i.e., "Male" or "Probable Male"). Seventy-six percent of the control participants concluded that the remains were "Male" or "Probable Male."

These results highlight the fact that exposure to extraneous context may bias forensic anthropological sex estimation conclusions—even when those conclusions are based on reliable, standardized measurements. These findings will inform the discipline's approach to the development of protocols, guidelines, and best practice standards to mitigate the effects of cognitive bias.

**Reference(s):**

1. Hartley, S. *A Hierarchy of Expert Performance as Applied to Forensic Anthropology.* Unpublished master's thesis; University of West Florida, 2020.
2. Langley N.R., L.M. Jantz, S.D. Ousley, R.L. Jantz, and G. Milner. *Data Collection Procedures for Forensic Skeletal Material 2.0.* University of Tennessee Knoxville, Tennessee, 2016.
3. France D.L. Observation and Metric Analysis of Sex in the Skeleton. In *Forensic Osteology: Advances in the Identification of Human Remains*, edited by K. Reichs, 163-186. Charles C. Thomas, Publisher, Ltd., 1998.

**Cognitive Bias, Metric Sex Estimation, Forensic Anthropology**



*Presenting Author