Journal of Archaeological Science 95 (2018) 33–39

Contents lists available at ScienceDirect

Journal of Archaeological Science

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

Zooarchaeology in the era of big data: Contending with interanalyst variation and best practices for contextualizing data for informed reuse

Hannah Lau[a,*,1], Sarah Whitcher Kansa[b]

[a] Interdepartmental Archaeology Program, University of California, Los Angeles A210 Fowler Building, Los Angeles, CA 90095, USA
[b] Alexandria Archive Institute, 125 El Verano Way, San Francisco, CA 94127, USA

ARTICLE INFO

Keywords:
Zooarchaeology
Data sharing
Big data
Interanalyst variation
Data comparability

ABSTRACT

New digital publication technologies facilitate the publication of primary data and increase the ease with which archaeologists are able to share, combine, and synthesize large datasets. The research prospects that these technologies make possible are exciting, but they raise the issue of how comparable the original datasets really are. In this study we demonstrate an issue associated with many archaeological datasets: interanalyst variation. We conduct two independent analyses of one zooarchaeological assemblage and compare data. We consider the implications of the challenge interanalyst variation poses within projects and across projects. We then make recommendations for zooarchaeologists specifically, and for archaeologists more broadly, who are interested in publishing primary datasets in order to improve future understanding of these data and facilitate their reuse. These recommendations include specific guidance of what information needs to be published along with primary datasets to facilitate their responsible reuse in other projects, recommendations for incorporating interanalyst variation studies into research programs, and suggestions about what to do should analysts discover systematic biases in their analyses stemming from interanalyst variation.

1. Introduction

New digital technologies and the ability to share, combine, and synthesize large datasets offer exciting opportunities for innovation in archaeological research. Archaeology has begun to embrace the new research possibilities made possible by "big data" (e.g. Borgman, 2015; Cooper and Green, 2016) and the normalization of data-sharing and primary data publication has opened up new possibilities for large synthetic analyses (Faniel and Yakel, 2017; Kansa and Kansa, 2011; Kansa and Kansa, 2014; Kintigh, 2006). Zooarchaeology as a sub-discipline within archaeology has been at the forefront of embracing the ability for digital technology, data-sharing, and open data publication (Kansa et al., 2014; McKechnie et al. 2015). For example, meta-analysis has facilitated zooarchaeological research tracing the inception and spread of domesticated animals (Arbuckle et al. 2014, 2016; Atici et al., 2017; Conolly et al., 2011). Other examples of the use of meta-analysis of zooarchaeological data include research on differing consumption of wild animals in the American Southwest (Kintigh et al., 2018; Spielmann and Kintigh, 2011), and cultural and environmental influences on native communities' choices in fisheries in the Pacific Northwest (McKechnie and Moss, 2016). These projects, and others like them, illustrate the utility of combining primary datasets, often from multiple different projects and analysts, in tackling major research questions at a regional scale.

The increasing adoption of primary data publication and open access publication makes large comparative projects more feasible and invites non-specialists to explore these datasets. But while it can be tempting to combine and compare these more easily-accessible datasets, issues of comparability and interanalyst variation must be addressed (Atici et al., 2013; Domínguez-Rodrigo, 2012; Huggett, 2015; Kansa and Kansa, 2011; Kansa et al., 2014; Kintigh, 2006). This is not a new issue within zooarchaeology, or in archaeology more broadly, but new technologies have increased the ease and scale of comparative work. As the process and practice of data publication become more normalized and widespread, it would behoove us to find solutions to this problem before filling digital archives with poorly-described data. Further, the rise of open access data sharing means that the consumers of zooarchaeological data may include people without zooarchaeological training, and who are less sensitive to how analytical variation can impact interpretation. This is true of consumers of any primary data produced by sub-field specialists in archaeology.

Archaeologists working with all types of nominal data derived from

* Corresponding author.
E-mail address: hannah.lau@ucla.edu (H. Lau).
[1] Present Address: Anadolu Medeniyetleri Araştırma Merkezi, Koç Üniversitesi, Istiklal Caddesi No. 181 Merkez Han, Beyoğlu, 34433, Istanbul, Turkey.

https://doi.org/10.1016/j.jas.2018.03.011
Received 27 July 2017; Received in revised form 7 March 2018; Accepted 31 March 2018
0305-4403/

observer classification have had to grapple with internalyst variability (e.g., for stone artifacts see Gnaden and Holdaway, 2000; for site types see Sadr, 2016). Primary zooarchaeological data, while based off observed phenomena that can be compared to known specimens, are ultimately subjective (Wolverton, 2013). The use of identification methods following the classificatory schema borrowed from zoology gives the appearance that zooarchaeological data can be taken at face value. There is, in fact, one correct answer as to which species an archaeological specimen comes from, but the evidence may not be sufficient for all analysts to accurately and confidently identify it. The degree to which different zooarchaeologists will be willing and able to identify specific specimens to particular levels of taxonomic specificity will vary by training, experience, access to resources for comparison (i.e. comparative specimens, images, genetic characterization), and time constraints for analysis. Zooarchaeologists are sensitive to these issues, and several have suggested methods to remedy issues associated with interanalyst variation in identification (Driver, 2011; Gobalet, 2001; Wolverton, 2013) and interpretation (e.g. Atici et al., 2013; Jones and Gabe, 2015).

Several zooarchaeologists have suggested best practices to permit data-consumers to assess the quality of zooarchaeological data (Driver, 2011; Gobalet, 2001; Kansa et al., 2018; Lyman 2002; Wolverton 2013), which is an important step towards ensuring comparability among datasets. Driver (2011) has emphasized the importance of transparency in identification methodology. Analysts should define their methods of identification prior to the start of analysis, assess each fragment individually, and explicitly list what taxa are most likely present within the assemblage, which taxa will be difficult to differentiate from one another and how the analyst will address these challenges. Gobalet (2001) echoes Driver in a call for transparency in methodology following his 2001 study of inter-analyst variation of fish bones from an archaeological site in California. He also recommends consulting people with deep and specific knowledge of fauna in a region and engaging multiple analysts as checks on bias throughout the process. Wolverton (2013) emphasizes the necessity of addressing the quality of zooarchaeological reports to give data consumers confidence in the results of analyses. This can be done, he argues, by directly addressing “Quality Control” and “Quality Assurance.” Wolverton identifies Driver's recommendations for best practices in zooarchaeological analyses as essential for “Quality Control” in zooarchaeological research. He suggests “Quality Assurance,” can be achieved through reporting diagnostic characteristics, assessing biometric data, “verification and reanalysis of random samples,” and the use of biochemical analyses such as aDNA and paleoresidues (2012:388–92). Increased precision of analyses offered by the expansion of work like ZooMS (e.g. Buckley et al., 2010; Welker et al., 2015) promises to make this more feasible.

Scholars have used blind tests to assess taphonomic and analytic bias in zooarchaeological assemblages. Some scholars have employed these types of experimental analyses to determine whether analysts can accurately identify bone surface modification by anthropogenic and non-anthropogenic forces (Blumenschine et al., 1996; Nims and Butler, 2017). Others have examined the replicability of zooarchaeological data between analysts and their effect on common measures of species and body part abundance employed in zooarchaeological research: Number of Identified Specimens (NISP), Minimum Number of Individuals (MNI), and Minimum Number of Elements (MNE) (e.g. Domínguez-Rodrigo, 2012; Morin et al., 2017). Nims and Butler advocate for incorporating blind reassessments into analyses of assemblages at different stages of the research project to identify changes in identification practices that may occur due to “protocol drift” (2017: 751) over the course of analysis. They also emphasize the importance of transparent reporting of analytic protocols, in line with Driver's and Wolverton's recommendations.

These recommendations are aimed at addressing the greater issue of trust in published data sets, which is, again, not a new issue in archaeology, but one that scholars are grappling with on a much greater scale with the increasing availability of published digitized datasets. In their study of data use and reuse in the social sciences, archaeology, and zoology Faniel and Yakel found that for archaeologists the two most important factors used to judge the trustworthiness of the dataset for reuse were the reputation of the data producer and the completeness of documentation of the dataset (2017:111). These two factors are widely used to assess zooarchaeological data among specialists (Driver, 2011; Gobalet, 2001; Wolverton, 2013). However, because they may be more difficult for non-specialists to evaluate, as more zooarchaeological data is made digitally available and to a broader audience, we need additional measures for users to gauge the trustworthiness of a given dataset. Here, we advocate the inclusion of an interanalyst variation study to ensure the quality of a dataset.

In this paper we demonstrate the benefits and challenges associated with interanalyst variation in the best of circumstances — where two researchers observe the same assemblage, under the same conditions, and with access to the same resources. Our explicit focus in this study is on the generation and comparability of primary data. As such we have chosen not to extend the study to examine how these differences affect derived zooarchaeological measures, such as MNI and MNE, although we acknowledge that NISP is, itself, a constructed metric. Studies by Atici et al. (2013), Morin et al. (2017), and others have shown that analysts may draw different interpretive conclusions from the same dataset based on different choices made in the construction of such metrics. Our concern here is with the primary data upon which such metrics are based. We then consider the implications posed by interanalyst variation within and across projects. Finally, we make recommendations for zooarchaeologists interested in publishing primary datasets in order to improve understanding of datasets and facilitate their reuse. These recommendations include being transparent about methodologies, incorporating interanalyst variation studies, where feasible, and what to do should analysts discover systematic biases in their analyses.

2. Material and methods

Both analysts examined the same assemblage of 468 specimens. This is a hand-picked assemblage from the Late Neolithic site of Domuztepe (ca. 6000-5450 BCE) in southeastern Turkey. The site was excavated under the direction of Elizabeth Carter (University of California, Los Angeles) and Stuart Campbell (University of Manchester) between 1995 and 2011 (Campbell et al., 1999; Carter et al., 2003). Both analysts are experienced zooarchaeologists who have worked primarily on collections from the Greater Ancient Near East. Both have analyzed large portions of the assemblages at Domuztepe (Campbell et al. 1999, 2014; Carter et al., 2003; Kansa et al., 2009a; Kansa et al., 2009b; Lau, 2016). While we undertook this exercise in order to align our work on faunal assemblages at Domuztepe, what we learned from the experience has broader applications across zooarchaeology and important implications for any aspect of primary data publication and comparison in archaeological research that requires classification in data production.

This study was analogous to instrument calibration in scientific analysis; in this case, however, the instrument was human. Therefore, we attempted to limit the number of possible variables that might affect our identifications. The bones were cleaned with water and air-dried prior to identification. Both analysts undertook their work at UCLA's Cotsen Institute of Archaeology in the same laboratory space to ensure the same lighting. Both analysts had access to the comparative collection at the UCLA Cotsen Institute of Archaeology Zooarchaeology Lab, which includes several specimens of each of the domesticated taxa found in this assemblage, and some of the wild fauna. In addition to the comparative specimens both analysts had access to the same guides and images (Boessneck, 1970; Helmer and Rocheteau, 1994; Hillson, 2005; Pales and Lambert, 1971; Prummel, 1988; Prummel and Frisch, 1986; Schmid, 1972; Walker, 1985; Zeder and Lapham, 2010), and 3-D