ORIGINAL PAPER

A method for the taphonomic assessment of bone tools using 3D surface texture analysis of bone microtopography

Naomi L. Martisius [1,2] • Shannon P. McPherron [2] • Ellen Schulz-Kornas [2,3,4] • Marie Soressi [5] • Teresa E. Steele [1,2]

Abstract

Increasingly researchers have employed confocal microscopy and 3D surface texture analysis to assess bone surface modifications in an effort to understand ancient behavior. However, quantitative comparisons between the surfaces of purported archaeological bone tools and experimentally manufactured and used bones are complicated by taphonomic processes affecting ancient bone. Nonetheless, it may be reasonable to assume that bones within the same deposits are altered similarly and thus these alterations are quantifiable. Here we show how unworked bones can be used to quantify the taphonomic effect on bone surfaces and how this effect can then be controlled for and incorporated into an analysis for evaluating the modified surfaces of purported bone tools. To assess the baseline taphonomy of Middle Paleolithic archaeological deposits associated with typologically identified bone artifacts, specifically *lissoirs*, we directly compare the surface textures of ancient and modern unworked ribs. We then compare the ancient unworked ribs and *lissoirs* to assess their differences and predict the ancient artifacts' original surface state using a multilevel multivariate Bayesian model. Our findings demonstrate that three of five tested surface texture parameters (*Sa*, *Spc*, and *IsT*) are useful for distinguishing surface type. Our model predictions show that *lissoirs* tend to be less rough, have more rounded surface peaks, and exhibit more directionally oriented surfaces. These characteristics are likely due to anthropogenic modifications and would have been more pronounced at deposition. Quantifying taphonomic alterations moves us one step closer to accurately assessing how bone artifacts were made and used in the ancient past.

✉ Naomi L. Martisius
nlmartisius@ucdavis.edu

[1] Department of Anthropology, University of California, Davis, Davis, CA, USA

[2] Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

[3] Department of Cariology, Endodontology and Periodontology, University of Leipzig, Leipzig, Germany

[4] Max Planck Weizmann Center for Integrative Archaeology and Anthropology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

[5] Faculty of Archaeology, Leiden University, Leiden, The Netherlands

Introduction

Archaeologists analyze and interpret artifacts of ancient material culture to learn about past human behaviors. A particular focus of the study of ancient material culture is to tease out the nuances of an object's production process, use, disposal, and post-depositional alterations. Microscopic alterations of an artifact's surface provide insight into such processes. Attempts at understanding the microscopic traces left on an artifact during its use life date back to the mid-twentieth century or even earlier (Curwen 1930; Semenov 1964; Vaughan 1981). Methodological developments continued in this field through the last half of the twentieth century, though most approaches remained qualitative in nature and therefore subject to discrepancies between researchers (Stemp et al. 2016). The need for quantitative development was recognized early on (Beyries et al. 1988; Grace 1989; Grace et al. 1985) but did not begin to take hold until the twenty-first century (Evans and Donahue 2008; Stemp and Stemp 2001, 2003). While various 3D quantitative methods have been employed (Stemp et al. 2016), the

use of confocal microscopy and ISO 25178 surface texture parameters (International Organization for Standardization 2012) has been more widely applied in recent years (Calandra et al. 2019a; Calandra et al. 2019b; Evans and Macdonald 2011; Galland et al. 2019; Henshilwood et al. 2018; Ibáñez et al. 2018; Macdonald et al. 2018; Macdonald et al. 2019; Martisius et al. 2018; Rosso et al. 2017).

The vast majority of surface texture analyses on archaeological artifacts have been devoted to understanding stone tools, which has led to substantial methodological development in this area (Evans and Donahue 2008; Evans and Macdonald 2011; Giusca et al. 2012; Ibáñez et al. 2018; Macdonald et al. 2018; Macdonald et al. 2019; Stemp and Chung 2011; Stemp et al. 2015; Stevens et al. 2010; Werner 2018). Even so, these methods have only infrequently been directly applied to the archaeological record (Caux et al. 2018; Galland et al. 2019; Ibáñez et al. 2014; Ibáñez et al. 2016; Stevens et al. 2010). Largely, this is due to the uncertainty of post-depositional effects such as striations and rounding of artifact surfaces as well as those with patinas (Caux et al. 2018; Galland et al. 2019; Werner 2018). There has been little methodological development devoted to the study of surface textures on other materials such as ochre (Rosso et al. 2017) or bone (Bradfield 2020; d'Errico and Backwell 2009; Lesnik 2011; Martisius et al. 2018; Watson and Gleason 2016), and even more rarely on taphonomic alterations of bone (Vietti 2016). Other 3D quantitative techniques applied to bone surface modifications such as cut marks, trampling traces, and tooth marks on archaeological materials have benefited from a fair amount of methodological advancement (Archer and Braun 2013; Bello and Soligo 2008; Bello et al. 2011; Gümrükçü and Pante 2018; Harris et al. 2017; Otárola-Castillo et al. 2018; Pante et al. 2017; Valtierra et al. 2020), which has been largely motivated by a debate over a report of cut-marked bones substantially older than 2.5 Ma (Dominguez-Rodrigo et al. 2010; McPherron et al. 2010). To ensure that interpretations of these controversial marks are secure, new methods including the use of Bayesian statistical analyses have been advanced (Harris et al. 2017; Otárola-Castillo et al. 2018). Employing Bayesian modeling with data derived from 3D models allows for probabilistic interpretations about the form of the bone modifications (Otárola-Castillo and Torquato 2018). However, there is currently no codified set of procedures for assessing the surface texture of archaeological bone microtopography. This study aims to begin this process and advance our ability to interpret ancient bone artifacts modified through post-depositional effects.

As current 3D quantitative methods stand, there are few baselines against which functional interpretations of bone microtopography can be made (d'Errico and Backwell 2009; Lesnik 2011; Martisius et al. 2018; Watson and Gleason 2016). In order to understand how and why a bone object was made and used, and directly relate modification to

behavior, certain comparative metrics must be in place. Especially when dealing with innovative quantitative technologies, it may not be enough to replicate a bone tool, use it, and then immediately apply those findings to the archaeological record. A number of post-depositional effects will have likely altered the ancient artifact to some extent, obscuring or perhaps imitating manufacturing and microwear traces. Traditionally, objects with clear taphonomic effects have either been excluded from such analyses or the entire surfaces were scrutinized to locate unaffected areas (Burroni et al. 2002). Though these qualitative initial assessments should always be undertaken, this may be a more difficult endeavor when studying objects made from organic components such as bone that have undergone chemical and physical alterations since initial deposition (Hedges 2002; Nielsen-Marsh and Hedges 2000).

Little is known about how the changes in bone geochemistry alter the microtopography of bone artifact surfaces (Krajcarz 2019; Orłowska 2018). There is evidence that some physical post-depositional processes partially affect surface texture on teeth (Böhm et al. 2019) and cut marks on bone (Gümrükçü and Pante 2018). In addition, it has been shown that bone surface roughness increases with bone weathering stages (Behrensmeyer 1978; Vietti 2016), but it is not yet clear how other post-depositional alterations may affect surface roughness and other surface texture parameters of bone microtopography including microwear. Resolving this is essential since many of the earliest bone technologies may preserve ambiguous microwear traces obscured by post-depositional effects. An essential component of the groundwork needed to understand these potentially ambiguous bone objects is the ability to quantitatively isolate taphonomic signatures. Assuming that each archaeological assemblage or deposit has its own measurable taphonomic effect due to diagenetic factors such as collagen loss, increased crystallinity, or dissolution (Hedges 2002; Nielsen-Marsh and Hedges 2000), precise measurements like those taken for surface texture analysis may be useful in this endeavor. Therefore, in this study, we show how unworked bones coming from the same archaeological deposit as purported bone tools can be used to quantify the taphonomic effect on bone surfaces and how this effect can then be controlled for and incorporated into an analysis for assessing modifications that occurred prior to the deposition of those purported bone tools. We do not aim to determine which diagenetic processes have affected the assemblages as chemical analyses would be required for an accurate assessment (Hedges 2002). This direct comparison should allow for the isolation and quantification of the taphonomic effects in the archaeological assemblages and the anthropogenic qualities of the worked bone in the sample, which will then provide the opportunity to predict the initial surface texture of the archaeological bone artifacts. By doing so, we aim to resolve some of the many confounding factors that can muddy the

interpretation of these artifacts. A further aim of this article is to establish an initial sampling protocol for applying 3D surface texture analysis to archaeological bones. The protocol will serve as a framework for the development of methods for the quantitative analysis of archaeological bone material.

Materials and methods

Bone tools, or technological objects made from an animal's skeletal remains, are commonly found in archaeological sites attributed to recent humans, i.e., after 50 ka (e.g., Bradfield et al. 2018; Hublin et al. 2020; Langley et al. 2016; Legrand 2007; Sidéra 1993; Tejero 2014; Tejero et al. 2016; Zhang et al. 2016; Zhang et al. 2018), and sometimes preserved in archaeological layers of other hominins (e.g., Backwell and d'Errico 2001; Gaudzinski 1999; Julien et al. 2015; Mania and Mania 2005; Radmilli and Boschian 1996; Sano et al. 2020; Soressi et al. 2013). Purported Neandertal-made formal bone tools, those manufactured with techniques specific to bone such as scraping and grinding (d'Errico et al. 2012), come from two late Middle Paleolithic sites in southwest France (Martisius et al. 2020; Soressi et al. 2013). Five rib fragments have been preserved in three different layers at Abri Peyrony and Pech-de-l'Azé I (Fig. 1a, b, c, d, and e) and are interpreted to be lissoirs, a French typological term meaning "smoothers" (Leroi-Gourhan 1968). All five of these objects exhibit morphological characteristics consistent with lissoir typology such as their elongated overall shape with a rounded or ogival distal end that exhibits polish (Averbouh 2000; Camps-Fabrer 1966; Lartet 1861; Mons and Stordeur 1977; Sonneville-Bordes 1960; Tartar 2009). We note that unless stated otherwise, here we use the term lissoir in this typological sense only

without the implied function in the same way that retouched flakes in the Middle Paleolithic are called scrapers.

Four of the five artifacts were taxonomically identified using nondestructive zooarchaeology by mass spectrometry, all of which were attributed to aurochs or bison (Martisius et al. 2020). The consistent use of large bovid ribs as the raw material for making these tools suggests they were strategically selected by Neandertals. The indications that these artifacts were shaped prior to use come from parallel striations along some of the artifact edges and surfaces, which are consistent with grinding against a coarse stone and scraping with a flint edge (Martisius 2019; Soressi et al. 2013). The initial publication describing four of these tools suggested they are the first specialized bone tools in Europe, due to their standardized shape and similarity to tools specifically made for working animal skins (Soressi et al. 2013). In addition, one of the artifacts preserves characteristics such as rounding of the upper reliefs and smoothing of furrows of the microtopography, features consistent with the use on a material like animal skin (Soressi et al. 2013). These types of bone tools are often found in later assemblages such as those from the Upper Paleolithic (Tartar 2009), and similar bone tools are even used by modern leather workers (Soressi et al. 2013). The presence of the same tool type in Neandertal contexts could suggest that Neandertal use of lissoirs was similar to humans during the subsequent Upper Paleolithic. Arguments for grouping these tools into a common category of use that spans Neandertal to modern times will require better quantitative data from the tools themselves.

Though qualitative observations are informative for describing bone surface modifications, a complementary quantitative approach will allow for the precise measurement of specific features of the bone microtopography and the

Fig. 1 Photographs of the lissoirs (a, b, c, d, e) and some of the unworked rib fragments (f, g, h, i, j, k, l) in this study at (a, g, i, j, k) Pech-de-l'Azé I (PA I) and (b, c, d, e, f, h, l) Abri Peyrony (AP); (a) PA I G8-1417, layer 4; (b) AP-4209, layer 3A; (c) AP-4493, layer 3B; (d) AP-10818, layer 3B; (e) AP-7839, layer 3B; (f) AP-6347, layer 3B; (g) PA I G8-1116b, layer 4; (h) AP-3468a, layer 3B; (i) PA I G8-1116a, layer 4; (j) PA I G8-1386b, layer 4; (k) PA I G8-1386a, layer 4; (l) AP-3468b, layer 3B. Photos of the unworked ribs courtesy of Chase Murphree. Images of the lissoirs adapted/modified from Martisius et al. (2020) and Soressi et al. (2013)
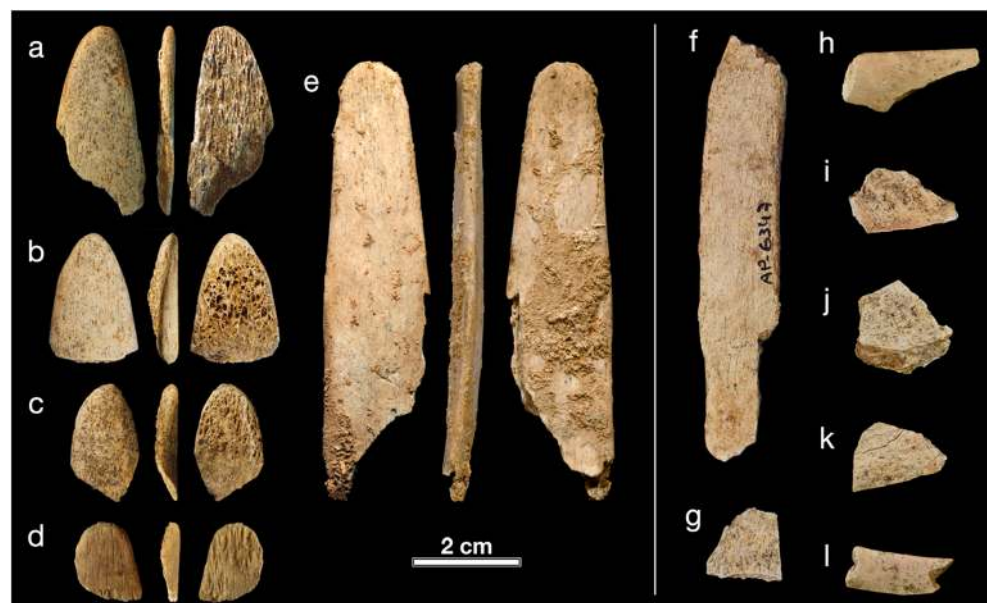
development of replicable standardized procedures. The aim of this study is to quantify the microtopographic differences between modern and ancient bone fragments and develop a method for predicting the initial surface texture states of ancient bone artifacts. Specifically, we compare modern unworked ribs and ancient unworked rib fragments preserved in the Middle Paleolithic archaeological layers associated with the *lissoirs* at Abri Peyrony and Pech-de-l'Azé I (Fig. 1f, g, h, i, j, k, and l). We expect the modern and ancient unworked rib fragments to exhibit quantitative differences indicative of 50,000 years of post-depositional microscopic alterations to the bone microstructure. The amount of surface texture differences obtained in this assessment will allow us to estimate the background taphonomy of the archaeological assemblages. Having the ability to quantitatively assess the changes that have occurred over millennia will be key in our ability to predict ancient unobserved surface textures for bone artifacts. The *lissoir* surfaces will then be compared with this taphonomic baseline to predict the range of their original surface textures. With the ability to quantify the taphonomic effect of ancient bone surfaces and utilize this effect to estimate the initial surfaces of purported bone tools, we aim to develop a method that can ultimately be used to assess the manufacture and use traces of archaeological bone artifacts. The study sample is systematically targeted based on qualitative observations of the different bone surfaces. We expect that observable microscopic distinctions (e.g., striations and smoothed surface features) between the *lissoirs* and the unworked rib fragments, both ancient and modern, should correspond to measurable quantitative differences. In fact, it would be surprising to find no distinctions between the surface types. This study is not meant as a replacement of traditional microwear analyses, but as a proof of concept to demonstrate the usefulness of quantifying the differences between bone surfaces. In conducting this analysis on both ancient and modern unworked bones to ones that are proposed to be manufactured and used as tools in the past, we begin to develop a framework for employing surface texture parameters in a meaningful way. With a more complete understanding of these parameters and how they relate to the various surface types, we can begin to formulate theoretical frameworks for applying more targeted analyses such as those aimed at distinguishing material wear as well as investigations into other bone objects that are even more challenging to analyze.

Samples

Five bone tools typed as *lissoirs* and 20 unworked rib fragments from the two Middle Paleolithic sites in France, Pech-de-l'Azé I and Abri Peyrony, were used in this study (Fig. 1; Table 1; Table S1). We targeted ungulate rib fragments preserved within the same archaeological layers and located in close proximity to the *lissoirs* knowing that taphonomy can

Table 1 Number of specimens used in this study by type including the site and layer in which they were preserved

Source	Layer	Specimen type	Number of pieces
Modern sample		Unworked rib fragment	8
Abri Peyrony	L-3A	*Lissoir*	1
		Unworked rib fragment	2
	L-3B	*Lissoir*	3
		Unworked rib fragment	11
Pech-de-l'Azé I	4	*Lissoir*	1
		Unworked rib fragment	6
	II-4	Unworked rib fragment	1
Total			33

vary across archaeological deposits. In one case, we were able to include an unworked rib fragment that was found in the same bucket of screened material as one of the artifacts (*lissoir*: AP-4493, ancient unworked rib: AP-4493b; Table S1). Recently, it has been shown that ungulate ribs exhibit similar surface roughness values (Vietti 2016), so comparisons including differing ungulate species should not bias results. The partially fresh modern cow (*Bos taurus*) ribs obtained from Archeoshop and scanned for a previous experimental study were used for comparative purposes (Martisius et al. 2018). These scans originated from surface molds taken from the flattest surfaces of eight unworked rib fragments at time 0 of the experiment. The *lissoirs* were scanned directly, while the unworked ribs had surface molds taken (President Jet Plus Light Body, Coltène, Altstätten, Switzerland), which were brought to the Max Planck Institute for Evolutionary Anthropology (MPI-EVA), Department of Human Evolution in Leipzig, Germany, for further analysis (Table S1; Table S2). Because the modern and ancient unworked ribs' scans were directly made on surface molds, comparisons between these two sample types will not be confounded by the type of material scanned, if these differences are significant at all. Previous research has shown that low-viscosity molding material adequately replicates microscopic surfaces for most quantitative measurements (Goodall et al. 2015; Martisius et al. 2018), so direct comparisons between the bones and molds should be reasonable. However, a recent study on miniaturized technological components showed molding material to underestimate the original surface (Baruffi et al. 2017). Though little research on quantifying replicated bone surfaces has been performed, our previous study found that molding material may not accurately replicate the shape of the peaks on the bone, resulting in decreased values for one surface texture parameter, arithmetic mean peak curvature [Spc] (Martisius et al. 2018). Because of this, we use Spc on molded bone surfaces with caution. Other surface texture parameters did not seem to differ appreciably between measurements on the

original bone and those on molding material (Martisius et al. 2018). Further, inconsistencies may arise at higher magnifications or due to human error in the molding process (Baruffi et al. 2017; Macdonald et al. 2018; Mihlbachler et al. 2019). Given that our analysis is conducted using a × 20 lens, molding inconsistencies are less likely to affect our results. In addition, each rib fragment had multiple molds taken, so their subsequent comparison was undertaken prior to scanning.

3D surface texture analysis

The Middle Paleolithic *lissoirs* and molds of the unworked ribs (modern and ancient) were scanned at multiple loci with a confocal disk-scanning microscope (μsurf mobile, Nanofocus AG, Oberhausen, Germany), using a × 20 lens (numerical aperture = 0.4, field of view = 0.8 mm^2) provided by the MPI-EVA in Leipzig, Germany. An attempt was made to scan the distal ends of the *lissoirs* in a systematic way using a crosswise formation with one scan in the center, two along the longitudinal axis, and two along the horizontal axis (Fig. S1). However, differential preservation did not allow for systematic sampling to this degree, so other scanning locations were incorporated using a coding system developed to identify where each scan originated (Fig. S1). For many of the artifacts, differential preservation constrained our ability to take multiple scans of the distal ends, so sampling was expanded to other areas of the tool. In doing this, we intend to quantify an overall "*lissoir* effect," which may include surfaces modified by manufacturing, use, or handling, rather than subject our statistical analysis to the stochastic effects of the small sample restricted to the distal end. All artifacts were oriented so that the longitudinal axis of the artifact was along the X-axis of the μsurf mobile with the distal end (tip) positioned at the extent of the X-axis. The distal end of the molds of the unworked rib fragments was arbitrarily assigned. From this point, a grid was set up from a central (C) position, most often 3.375 mm proximally from the tip. Given the field of view, each possible scan in the grid was 0.8 mm^2 with a minimal amount of overlap in each area (Fig. S1). After each scan was taken, scan quality was reviewed. Those with 95% or more of the surface points measured were accepted for further study (Martisius et al. 2018; Schulz et al. 2010, 2013a). Scans of lesser accuracy were remeasured altering pitch, gain, exposure, or brightness values until 95% of the microtopography was captured.

The scans were then transferred to the MountainsMap Premium v. 7.4.8076 Analysis software (Digital Surf, Besançon, France) for 3D surface texture analysis. To construct meshed axiomatic 3D models of the modern and archaeological pieces (Fig. 2), the following procedures were applied: *leveling* (least square method), *mirroring* the y- and z-axes for surface molds, and *removal of outliers* (isolated and edge outlier removal, with normal strength, and hole fill in

< 225 points, noise removal), *non-measured points fill in* (smoothing method), and *form removal* using a polynomial of 2nd degree (Martisius et al. 2018; Schulz et al. 2010, 2013a; Schulz et al. 2013b).

Before the meshed axiomatic 3D models of the artifacts could be analyzed, an additional in-house step was taken to ensure data would be informative. All models and corresponding digital images were visually inspected for indications of sizable taphonomic alterations that removed the cortical bone surface and exposed internal structures (Fig. S2). Surfaces with large craters or a number of exposed pores were excluded from further analyses because measurements of these surfaces would be skewed towards capturing the variation of the bone structure and not the effects indicative of anthropogenic alterations (Martisius et al. 2018). Similarly, modern rib scans with large irregular trenches that may have been formed during periosteum removal were excluded from further analyses (Fig. S2). Only the meshed axiomatic 3D models with at least 90% of the cortical surface preserved were kept for statistical modeling. Consequently, many of the surface scans did not meet this threshold and only a small number of meshed axiomatic 3D models of each specimen were kept for analysis (Table S1; Table S2). Models from both ancient and modern unworked ribs were included in similar proportions (modern unworked ribs, 68%; ancient unworked ribs, 67%), while models of *lissoir* surfaces were disproportionately affected by large irregular surface features thus included at a rate of only 8%. Figure 3 shows an example of one of the artifacts and its surface locations used for statistical modeling.

Statistical modeling

We chose five ISO 25178 parameters for statistical modeling: arithmetic mean height [*Sa*], autocorrelation length [*Sal*], arithmetic mean peak curvature [*Spc*], upper material ratio [*Smr1*], and isotropy [*IsT*] (Fig. 4; International Organization for Standardization 2012). The first four ISO 25178 parameters were selected because in a prior experimental study they showed some degree of differentiation between variously modified bone surfaces (Martisius et al. 2018). The fifth parameter, *IsT*, was also included because it should be an indicator of the anthropogenic traces on the *lissoirs* given that their surfaces retain nonrandomly oriented striations from manufacture and use.

Statistical modeling follows protocols in Martisius et al. (2018). The five ISO 25178 parameters were log transformed to stabilize variances and distributions. The statistical model for the observations *Y*, a matrix of *p* = 5 columns (log-transformed ISO 25178 parameters), and *n* = 136 rows (3D models) is a multivariate mixed model of the form:

$$Y = XB + ZU + E;$$

Modern unworked ribs
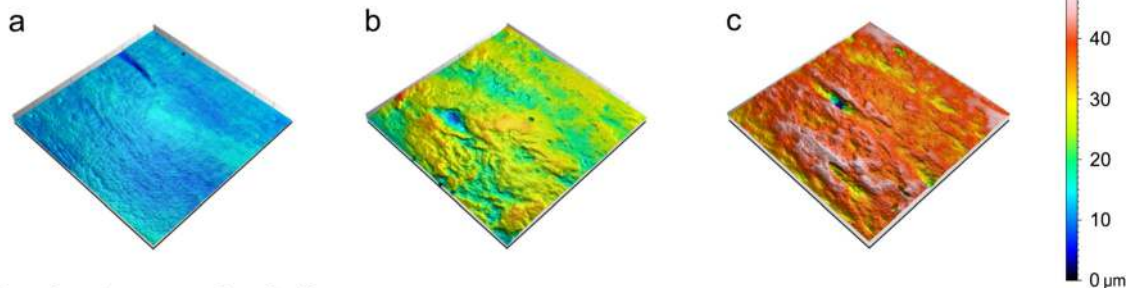
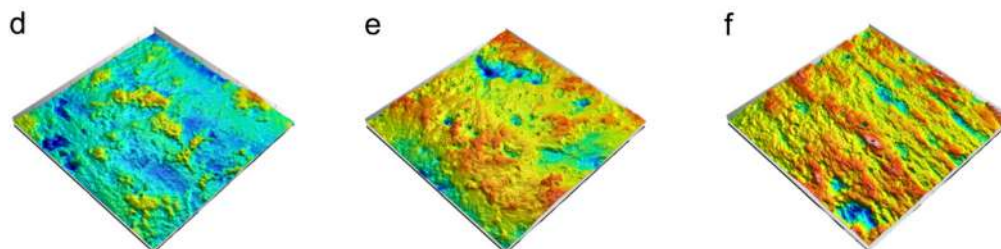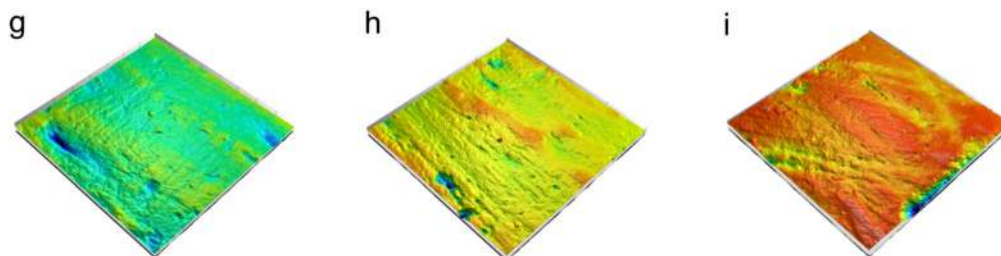## Ancient unworked ribs

Ancient *lissoirs*

Fig. 2 Meshed axiomatic 3D models (0.8 mm²) of modern rib fragments (a, b, c), ancient rib fragments (d, e, f), and ancient *lissoirs* (g, h, i) at Abri Peyrony (d, g, h, i) and Pech-de-l'Azé I (e, f); (a) VII3; (b) VII1; (c) VI4; (d) AP-2166, layer 3A; (e) PA I G8-1386, layer 4; (f) AP-3468, layer 3B; (g) AP-7839, layer 3B; (h) AP-4209, layer 3A; (i) AP-10818, layer 3B. Color of surfaces correspond to height on the z-axis in upper right-hand corner

(Amemiya 1994) where XB represents the fixed effects, ZU the random effects, and E the residual error. Here, ZU captures idiosyncratic bone effects of each *specimen* and the *source* from which they were acquired. This includes the layer in which the archaeological specimens were preserved. U is a 38×5 matrix of random intercepts, where each column of U contains 33 unique *specimen* effects plus 5 unique *source* effects. The five columns of U correspond to the five surface texture parameters. Z is a 136×38 matrix of ones and zeros, indicating the *specimen* and *source* membership of each surface scan.
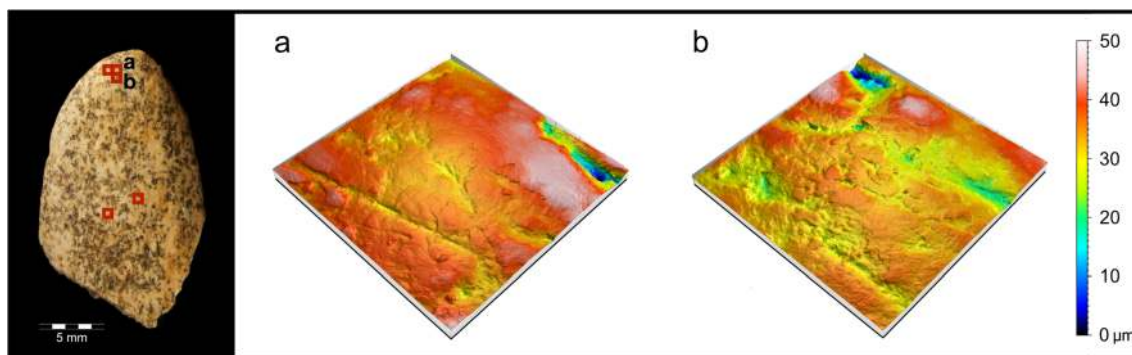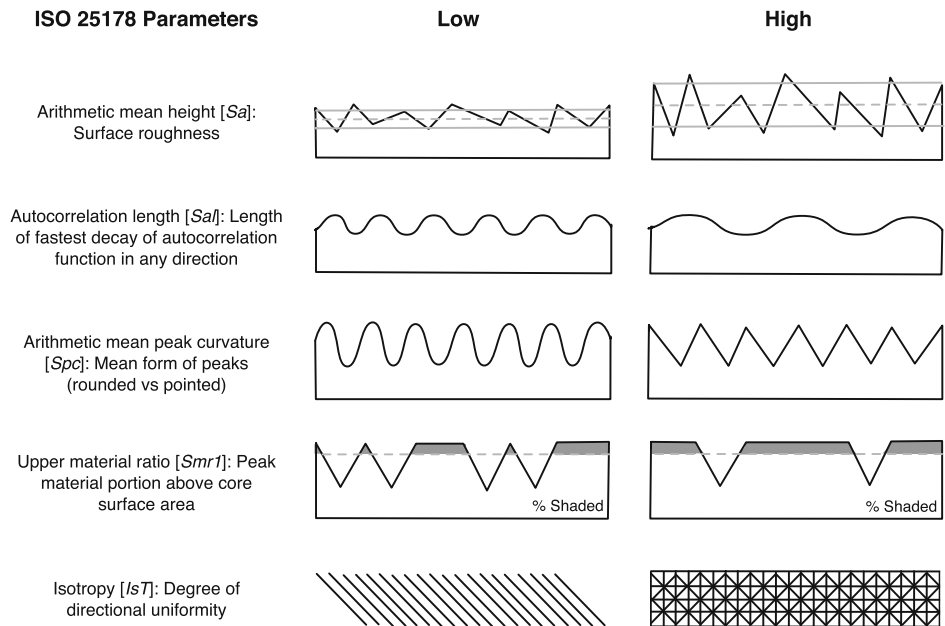

Fig. 3 Photograph and (a, b) meshed axiomatic 3D models of AP-4493 from Abri Peyrony. Red boxes indicate locations of scans used for statistical modeling. All scanned areas are 0.8 mm². Image of the artifact adapted/modified from Soressi et al. (2013)

Fig. 4 2D depictions of ISO 25178 3D texture parameters. Illustrations on the left show examples of surfaces with low values, while those on the right have high values (adapted/modified from Kaiser et al. (2016) and Martisius et al. (2018))

ISO 25178 Parameters	Low	High
Arithmetic mean height [Sa]: Surface roughness		
Autocorrelation length [Sal]: Length of fastest decay of autocorrelation function in any direction		
Arithmetic mean peak curvature [Spc]: Mean form of peaks (rounded vs pointed)		
Upper material ratio [Smr1]: Peak material portion above core surface area	% Shaded	% Shaded
Isotropy [IsT]: Degree of directional uniformity		

XB are the fixed effects in the model. After fitting two models of increasing complexity with different effects, we compared widely applicable information criterion (WAIC) scores (Vehtari et al. 2017) and found that a design (M1) with the single fixed effect *specimen type* (ancient *lissoir*, ancient unworked rib, and modern unworked rib) generated model predictions best (Table 2). For M1, *B* is a 3 × 5 matrix of fixed effects for each *specimen type,* for each surface texture parameter. The design matrix *X* is a 136 × 3 matrix of zeros and ones. *E* is a 136 × 5 residual matrix. Therefore, M1 is a multilevel, multivariate Bayesian model, including a fixed effect (*specimen type*), random effects (*specimen* and *source*), and *error* (Table 2).

Compared with the results in Martisius et al. (2018), variation in surface texture parameter values was over-dispersed with respect to the Gaussian distribution, so we chose the multivariate Student *t* distribution. The *specimen* and *source* random effects were adequately modeled by multivariate Gaussian distributions, based on goodness of fit checks. See Martisius et al. (2018) for further details of the multivariate mixed model.

We used a Hamiltonian Markov chain Monte Carlo method to estimate statistical model effects, using the library rstan version 2.19.3 (Stan Development Team 2019) of the statistical computing language R version 4.0.0 (R Core Team 2020). We allowed a 4000-iteration warm-up for all four chains and generated 2000 samples per chain for a total of 8000 posterior samples for inference. Scaled and squared Mahalanobis distances between observations and predicted values were examined to check goodness of fit. Theoretical quantiles of the F-distribution (Roth 2013) were used to compare Mahalanobis distances using a quantile-quantile plot (Fig. S3).

The posterior samples were used to predict the *lissoir* initial surface state for the five ISO 25178 parameters using

$$\Delta ISO = AR_{ISO} - MR_{ISO}$$

$$IL_{ISO} = AL_{ISO} - \Delta ISO$$

where ΔISO represents the contrast of the two unworked rib *specimen types* (*AR*: ancient unworked rib and *MR*: modern unworked rib) for each ISO 25178 parameter. *IL* is the predicted initial *lissoir* surface state and is calculated by subtracting each ΔISO from the ancient *lissoir* (*AL*) parameter values.

Results

The inclusion of *specimen* as a random effect allows the statistical model to take the idiosyncratic properties of individual bones into account. Similarly, *source* was included as a random effect, so model predictions incorporate any shared

Table 2 Effects included in each model, their representation in the design space, and WAIC score differences relative to the score for model M1

Model	Effects	Design matrix	Δ WAIC
M0	Random	Specimen + source + error	9
M1	1 fixed + random	Specimen type + specimen + source + error	0

effects of the specimens' origin including the archaeological deposits in which the bones were preserved. Because *specimen* and *source* were adjusted for in the statistical model, the model predictions should have captured the differences between the *specimen types*.

A pairwise scatterplot matrix of the five surface texture parameters displays all observations, and colored ellipses indicate 95% posterior probabilities for the estimated means of each parameter pair, for *specimen type* (ancient *lissoir*, ancient unworked rib, and modern unworked rib) and the predicted initial *lissoir* surface state (Fig. 5). Paired comparisons are useful for determining which surface texture parameters best distinguish the *specimen types*. There is some degree of overlap between most of the comparisons, but many of the parameter comparisons show clear trends distinguishing the ancient unworked rib fragments from both the

modern ribs and the *lissoirs*. Surprisingly, many of the comparisons do not distinguish the modern unworked ribs from the ancient *lissoirs*, though the *lissoir* values are often clustered in one area of the modern samples' total variation, values that tend to be intermediate between the modern and ancient unworked ribs (Fig. 5).

The modern and ancient unworked ribs show the clearest distinctions using the three parameters, *Sa*, *Spc*, and *IsT*, and in the comparison, *Spc:IsT*, the model predictions show no overlap between these two bone types (Fig. 5). Generally, the comparisons with *Sa*, *Spc*, and *IsT* indicate that the ancient unworked ribs have higher values than the modern unworked ribs. This demonstrates that the ancient samples are more rough, have more pointed surface peaks, and have more isotropic surfaces when compared with the modern unworked ribs. *Sal* shows a slight trend with the ancient samples having
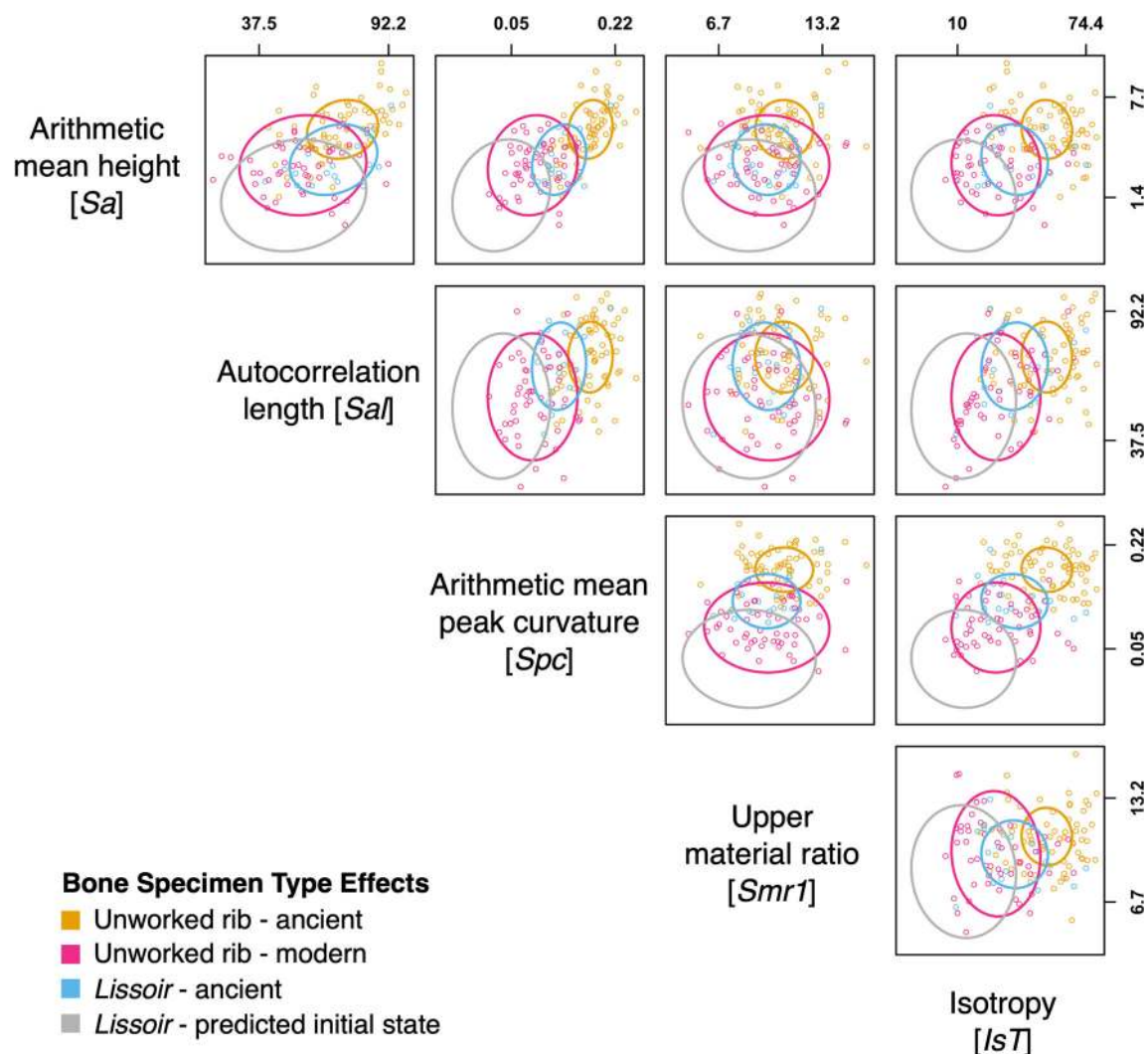

Bone Specimen Type Effects
- 🟧 Unworked rib - ancient
- 🟥 Unworked rib - modern
- 🟦 *Lissoir* - ancient
- ⬜ *Lissoir* - predicted initial state

Fig. 5 Scatterplot matrix of bone *specimen type* effects. Plots show all observations displayed in the pairwise space of parameters: surface roughness [*Sa*], autocorrelation length [*Sal*], peak curvature [*Spc*], upper material ratio [*Smr1*], and isotropy [*IsT*]. Observations are representative of bone *specimen type* effect (blue, ancient *lissoir*; yellow, ancient unworked rib; pink, modern unworked rib) and ellipses

are 95% credibility sets for the mean of each pair of parameters, made with ellipse version 0.3–8 (Murdoch et al. 2013). The gray ellipses represent the predicted initial *lissoir* state for each pair of parameters. Axes are on the log scale, but tick labels are in original measurement units and placed at the 5th and 95th percentiles

the highest values, meaning they tend to have lower frequency surfaces. There is no clear trend using *Smr1*, which indicates that the ancient and modern unworked ribs have similar portions of material in the peaks of their surfaces.

The model predictions indicate similar trends when focusing on the two ancient *specimen types* (*lissoirs* and ancient unworked ribs), though these distinctions are less prominent than those of the two unworked rib samples (Fig. 5). There is little overlap for three of the parameters, *Sa*, *Spc*, and *IsT*, while *Sal* and *Smr1* are less informative. In the three clearest comparisons (*Sa:Spc*, *Spc:IsT*, and *Sa:IsT*), the *lissoir* surfaces have the lowest values, while the ancient unworked rib fragments have higher values. This indicates that the *lissoir* surfaces are less rough, have more rounded surface peaks, and have more anisotropic surface textures as compared with the ancient unworked rib fragments. *Smr1* also shows a slight trend with *lissoirs* generally having lower values implying that they have a smaller portion of material in the peaks of the surface. *Sal* values are not well distinguishable between the ancient *specimen types* indicating that both types of artifact surfaces have similar wavelengths or frequencies.

The initial surface states of the *lissoirs* are predicted to have the lowest values for most of the surface texture parameters, but these predictions also have the largest variance due to the uncertainty in all aspects of the model used to make these predictions. When compared with the modern unworked ribs, a similar trend like what is observed for the two ancient *specimen types* (*lissoirs* and ancient unworked ribs) is observed. The model predicts that the initial *lissoir* surfaces would have been less rough, have more rounded surface peaks, and have more anisotropic surface textures compared with the unworked ribs deposited alongside them.

Discussion

Our results indicate that there are quantifiable differences between some of the *surface types* and that these differences have significant implications for directly applying surface texture data of experimentally worked bone to ancient artifacts. When doing so using 3D quantitative data, a correction factor representative of the background taphonomy will need to be employed in order to make accurate interpretations based on modern experimental data.

When analogous bones differ

Of importance are the quantitative differences between the ancient and modern unworked rib samples. Given that these two sample types are distinct in many of the pairwise comparisons (Fig. 5), it would not be wise to make interpretations based on a direct comparison of the surface texture values for modern and ancient bones. If we assume by analogy that

ungulate ribs in the past have a similar microstructure to ungulate bones today, then we would assume that any differences observed are likely the result of taphonomic alterations that have accumulated over 50,000 years (Soressi et al. 2013). Pinpointing the exact diagenetic alterations is beyond this study, because there are likely to be multiple correlated factors at play that are more easily detected via chemical analyses (Hedges 2002). Because of multiple possible taphonomic factors, the microtopography of the unworked ribs have become rougher with more pointed surface peaks over time (Fig. 6). A possible explanation could be related to the dissolution of the bone surface, which increases porosity and contributes to the loss of collagen (Hedges 2002). Additionally, the ancient surfaces tend to be more isotropic, which means that they are less directionally oriented. Bone is a naturally oriented and hierarchically organized structure (Fig. 2c; Rho et al. 1998); this is exhibited by the low *IsT* values of the modern unworked ribs (Fig. 5). Because the ancient unworked ribs have higher *IsT* values and therefore have less oriented surfaces, the taphonomic processes that affected the bone likely occurred in a fairly consistent manner over the bone surface obscuring the structure of the bone fibers.

Though the differences between the modern and unworked bone samples are likely the result of microscopic changes to the bone surfaces over millennia, there are a few other factors to consider. It is possible that one or both sample types are biased in some way. For example, the modern specimens derive from cow ribs, but it is not clear which ungulate species the ancient unworked rib fragments originate from. Due to their fragmentary nature, it is difficult to make species identifications of ancient ribs based on morphological characteristics (Martisius et al. 2020). For this study, we used ancient ungulate ribs because it is likely that the bone
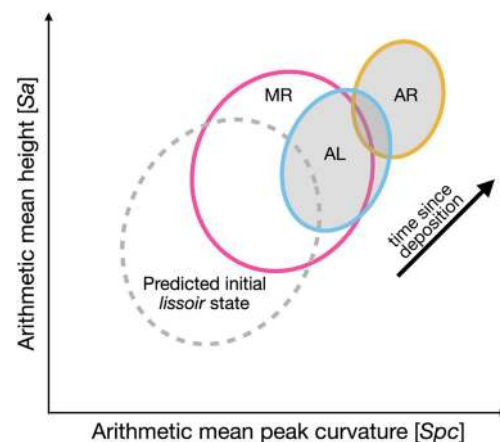

Fig. 6 Trajectory of surface texture parameters after millennia of taphonomic effects. Ellipses are the model predictions of the *specimen type* effects [blue, ancient *lissoir* (AL); yellow, ancient unworked rib (AR); pink, modern unworked rib (MR)] for surface roughness [*Sa*] and peak curvature [*Spc*]. The filled-in ellipses represent the bone samples with unknown taphonomic alterations. The dashed ellipse represents the hypothesized initial state of the *lissoirs* for the two displayed parameters

microtopography of differing ungulate species are very similar given their evolutionary history and common growth development (Currey 2006; Gifford-Gonzalez 2018). In addition, it has been demonstrated that the rib shafts of some ungulates (deer, elk, horse, and llama) have statistically similar initial surface roughness values (Vietti 2016), but other surface texture parameters have not been tested. Further, there may be variation from one animal to another including sex and age differences resulting in a broader range of variation; other factors including bone freshness may also bias the analysis and contribute to the patterning observed in this analysis (Karr and Outram 2015). Even so, the archaeological unworked rib sample should capture a reasonable amount of variation as it derives from 20 rib fragments from three different archaeological layers at two sites. The modern unworked ribs derive from one source, so it is possible that this sample does not include the full extent of modern cow rib variation. It should be noted that the statistical model indicates that the modern unworked rib sample has greater variation than the comparable ancient sample for most of the surface texture parameters (Fig. 5), which demonstrates that there is considerable variation within the one source. It is not clear if any of the aforementioned factors contributed to the differences in the two samples, so we can only speculate on the main causes for their distinctions. Whether the difference is a result of a biased sample or due to taphonomic alterations that accumulated over 50,000 years, we should be aware of the possible biases in our modern comparative specimens and directly compare surface texture of modern and ancient samples with caution (Karr and Outram 2015).

The effect of being a *lissoir*

Given that the ancient bone *specimen types* (ancient unworked ribs and *lissoirs*) originate from the same archaeological deposits and in close proximity to each other, it is likely that the quantifiable differences between these two samples are not the result of differing taphonomic alterations but of the anthropogenic qualities of the Neandertal-worked bone. This outcome is not surprising but expected given the nonrandom sampling strategy employed and the observed qualitative microscopic differences (Fig. 2; Fig. 3). Of importance is our ability to quantitatively describe these differences.

Sa, or surface roughness, is distinctly lower in the *lissoir* surfaces compared with the ancient unworked ribs (Fig. 5). This and other surface roughness parameters have been shown to distinguish unworked and worked surfaces in experimental samples, especially in those worked for long time intervals (Bradfield 2020; Martisius et al. 2018). In our previous study, we showed that *Sa* values decrease as a bone is worked over time (Martisius et al. 2018). Even in the examples where *Sa* was the highest at the beginning of the experiment due to modification using stone implements, the *Sa* values steadily

decreased and in some cases were lower than the original unworked bone surfaces as the experiment progressed (Martisius et al. 2018). Deep parallel striations along some of the edges and surfaces of the *lissoirs* are consistent with shaping using a stone implement (e.g., Fig. 2i), though it is not clear how extensively the bones were shaped prior to use (Martisius 2019; Soressi et al. 2013). The overall lower *Sa* values indicate that the *lissoirs* were likely modified through use, possibly after shaping, which led to the reduction of surface roughness.

Though most of the *lissoir* 3D models used in the statistical analysis are located within 1.5 cm of the distal end ($n = 14$), a small portion comes from other locations beyond the main working area of the tools ($n = 5$). These locations were added to increase the number of usable 3D models without irregular surface features such as sizable pores (Fig. S2) and reduce the stochastic effects of small sample size in the statistical model. These 3D models, including the ones in the active area of the tools, preserve a variety of empirically observed microwear traces including regularly oriented striations indicative of manufacturing and use (e.g., Fig. 2 g, h, and i). The inclusion of a variety of locations within the *lissoir* sample should reduce the *lissoir* effect for *Sa* and increase its variance. Specifically, *Sa* values would presumably be lower if only locations with visible use-related wear and without larger features resulting from likely contact with a stone implement were included (Fig. 2i), but the model predictions still indicate overall low values compared with the unworked bone (Fig. 5). The cause of the reduced *Sa* values is most likely from the use as a tool given that the majority of 3D models are located within the vicinity of the distal end, but other factors such as handling or rubbing inside a sack may have a similar effect and should be quantitatively examined. Nonetheless, the vast majority of the *lissoir* surfaces are less rough compared with the ancient unworked bone; if we restricted the analyses of the *lissoirs* to only the clearly worked surfaces, then we would expect the differences to be even more extreme.

Another surface texture parameter that indicates that the two ancient bone samples differ is *IsT* or isotropy (Fig. 5). Isotropic surfaces are not patterned with respect to direction (Fig. 4). This is in contrast to anisotropic surfaces that have directional properties. Though natural bone surfaces should show some degree of anisotropy due to its differently oriented and hierarchically organized structure (Fig. 2c, f; Rho et al. 1998), the ancient unworked rib surfaces (Fig. 2d, e, f) in this study are more uniform in their directional properties (more isotropic) than those of the *lissoirs* (Fig. 2 g, h, i; Fig. 3). The more extreme anisotropic properties of the *lissoirs* likely result from the striations and grooves produced during unidirectional manufacture and use. Our study indicates that these anthropogenic traces are more directionally patterned than the ancient unworked bone structures. Interestingly, some of the modern unworked bone surfaces are even more anisotropic

than the *lissoirs* (Fig. 5). It is possible that the same factors that caused the ancient unworked bones to be more isotropic also had the same effect on the *lissoirs*, partially obscuring the directionally oriented striations.

Spc, peak curvature, is the third surface texture parameter that indicates that the two ancient surface types differ (Fig. 5). *Spc* was included in the analysis with caution knowing that the dental molds used on the unworked bone surfaces (both ancient and modern) may lead to the reduction of their values due to possible replication inconsistencies of the shape of the bone peaks (Martisius et al. 2018). Even so, *Spc* is lower in the *lissoir* surfaces indicating that they have more curved surface peaks compared with the ancient unworked rib fragments. Further, if scans were made directly on the unworked bone surfaces and not on the dental impression material, this difference between the two surface types might be more pronounced. This pattern is expected, especially if the *lissoirs* were used on a soft, supple material such as animal skin as was originally determined for one of the Middle Paleolithic *lissoirs* (Soressi et al. 2013). Due to its material properties, animal skin should penetrate all levels of the bone microtopography rounding out the peaks of the surface. This pattern was demonstrated in Martisius et al. (2018), which showed that the experimental samples worked with leather and fresh hide had the lowest *Spc* values. Though other uses are likely to produce rounded peaks, the distinct values of the *lissoirs* compared with the unworked ribs indicate that the *lissoir* surfaces were most likely modified by a soft, supple material.

Though *Smr1* is less informative as there is considerable overlap in the values of this surface texture parameter (Fig. 5), there is some indication that *lissoir* surfaces are different compared with the ancient unworked rib surfaces. The *lissoir* *Smr1* values are slightly lower than the majority of ancient unworked bones, meaning less of the material is in the peaks of the *lissoir* surfaces. Significantly, all of the *Smr1* values of the ancient specimens are within the overall range of the modern specimens. This indicates that *Smr1* may not be useful for distinguishing bone surface microtopography.

Sal, or autocorrelation length, which describes the frequency of wavelength of the surface (Fig. 4), does not distinguish the two ancient bone types (Fig. 5). It does, however, indicate that there may be differences between the modern and ancient bones as the ancient specimens cluster at the higher range of the overall *Sal* values. This demonstrates that the ancient bone specimens generally have low frequency or long wavelength surfaces, while the modern bone tends to have greater diversity in this respect. This pattern could be explained through the dissolution of the bone surface, a chemical process that is more likely to cause the microscopic loss of bone material over the whole surface rather than in isolated areas (Hedges 2002).

Towards an understanding of ancient bone through surface texture

A surprising result of the analysis is that the surface texture parameters for the modern unworked ribs and the ancient *lissoirs* overlap (Fig. 5). Taking the statistical results at face value, one might assume that the purported *lissoirs* are not bone tools at all but unworked ribs. This line of reasoning would only work if the *lissoir* sample was not affected by the same taphonomic alterations as the ancient unworked ribs, a highly unlikely scenario. A more likely explanation is that the modern ribs and the *lissoirs* exhibit converging surface texture values but from different causes. Even though quantitative methods allow us to be more precise with our analyses, many of the surface texture parameters are still subject to equifinality and are thus statistically indistinguishable (Lyman 2004). These parameters are averaged values over the area of a measured surface, so a variety of surfaces may produce similar values (International Organization for Standardization 2012). For example, Figs. 7 and 8 demonstrate two different examples of one of each of the *specimen types* whose surface texture values cluster together in the pairwise space of parameters in Fig. 5. The 3D models from Fig. 7 cluster closely in the comparison with *Spc* and *Smr1*, while those in Fig. 8 cluster in the comparison with *Sa* and *Spc*. All values for each surface parameter are displayed indicating differences in the non-highlighted values. It is important to recognize the visual differences between each *specimen type*, yet they exhibit similar values for at least two of the surface texture parameters (Fig. 7; Fig. 8). Employing these parameters on bone microtopography is helpful for comparing and describing broad trends but the parameters and what they are measuring need to be completely understood when trying to discriminate smaller differences. To employ this type of analysis, a good understanding of the samples is required. This includes predictions about what values the surface texture parameters might show, as those parameters should not be blindly applied to a sample in search of differences or similarities. Given that each of the measured surface areas has one value per surface texture parameter, it is highly likely that variously affected surfaces could exhibit a similar value. Including one or two parameters in an analysis will result in the loss of the bigger picture. Employing multivariate analyses on this type of data should reduce the effects of equifinality. Similarly, a qualitative component should be incorporated into this type of analysis to obtain a fuller understanding of the analyzed surfaces.

Though surface texture does not always distinguish two differently modified bone surfaces (Martisius et al. 2018; Watson and Gleason 2016), especially when using a small number of parameters, identifying and utilizing the parameters that can help explain the differences in two bone samples are crucial. In this analysis, the differences between the ancient
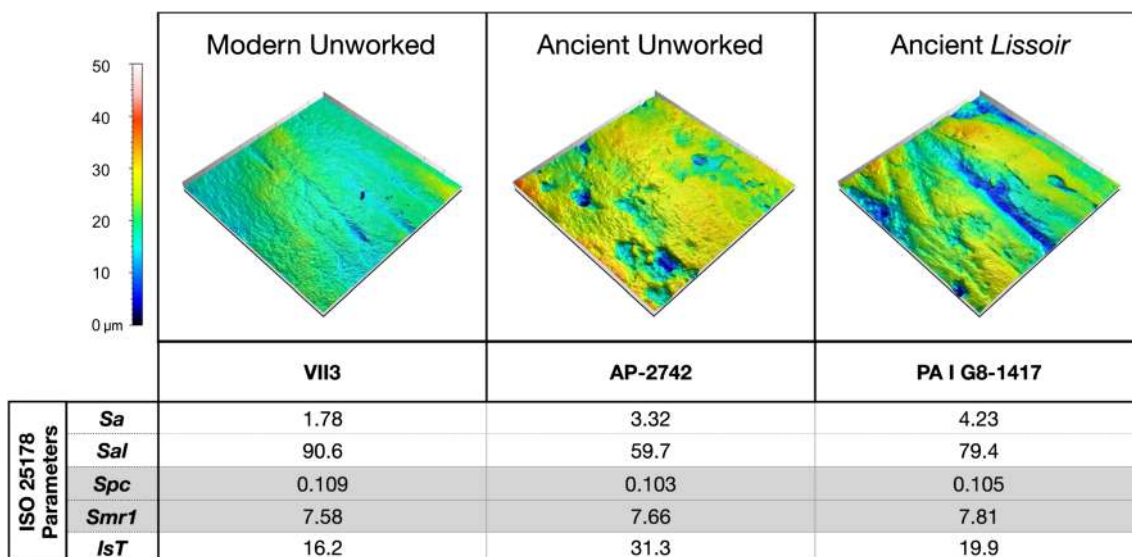
		Modern Unworked	Ancient Unworked	Ancient *Lissoir*
		VII3	AP-2742	PA I G8-1417
ISO 25178 Parameters	Sa	1.78	3.32	4.23
	Sal	90.6	59.7	79.4
	Spc	0.109	0.103	0.105
	Smr1	7.58	7.66	7.81
	IsT	16.2	31.3	19.9

Fig. 7 Meshed axiomatic 3D models ($0.8\ mm^2$) of one of each of the *specimen types* with the values for each of the surface texture parameters. The highlighted rows indicate values that are similar and cluster in the pairwise space of parameters (peak curvature [*Spc*] and upper material ratio [*Smr1*]) when plotted

and modern unworked bone samples can be used in an attempt to understand how taphonomy changes bone over time and what this might mean for assessing bone artifacts with a variety of surface alterations. Given that the archaeological bones in this study derive from sediments that are about 50 ka (Soressi et al. 2013), taphonomic alterations are the most likely cause of the differences between the ancient and modern unworked bone samples. Bone surface loss through dissolution is proportional to the external surface area (Hedges 2002), and because ribs are roughly of the same size, their alterations within the same deposits should be comparable. It is possible that taphonomy affects differently modified surfaces in

slightly different ways. For example, there is some evidence to suggest that compressed and polished bone surfaces resist the effects of weathering (Bradfield 2020; Moore 2013), which might apply to other taphonomic alterations, though this concept has not yet been tested on the surface texture of bone microtopography. Alternatively, exposing the internal structure of a bone during the manufacture and use of bone tools may accelerate diagenetic processes leading to bone surface microtopography that has been altered to a greater extent. The percentage of surface scans with sizable irregular features removed from this study could support this (*lissoirs* = 92%, ancient unworked ribs = 33%) (Fig. S2).
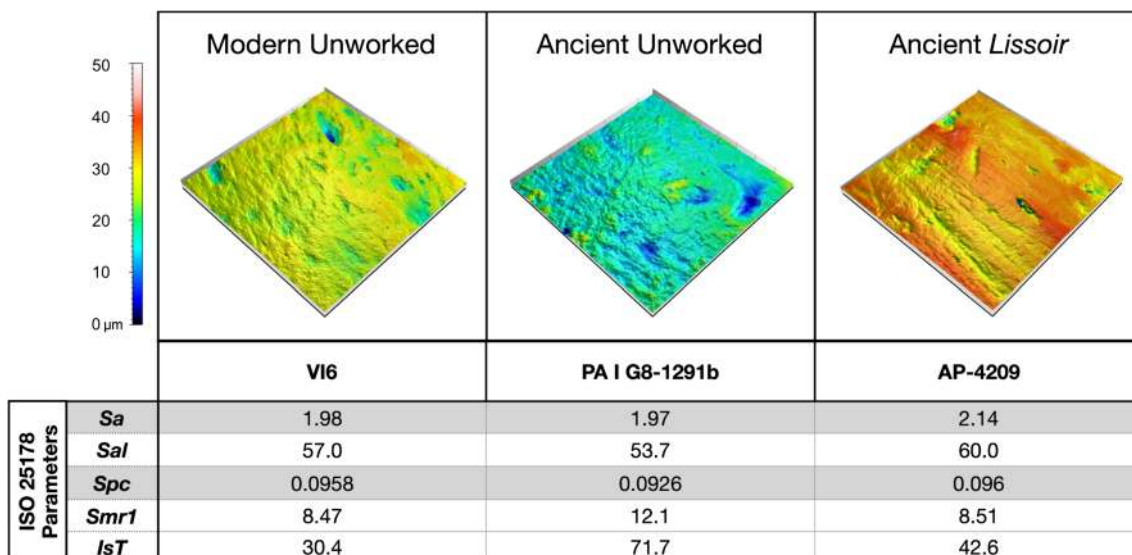

		Modern Unworked	Ancient Unworked	Ancient *Lissoir*
		VI6	PA I G8-1291b	AP-4209
ISO 25178 Parameters	Sa	1.98	1.97	2.14
	Sal	57.0	53.7	60.0
	Spc	0.0958	0.0926	0.096
	Smr1	8.47	12.1	8.51
	IsT	30.4	71.7	42.6

Fig. 8 Meshed axiomatic 3D models ($0.8\ mm^2$) of one of each of the *specimen types* with the values for each of the surface texture parameters. The highlighted rows indicate values that are similar and cluster in the pairwise space of parameters (surface roughness [*Sa*] and peak curvature [*Spc*]) when plotted

For the purposes of this study, we have estimated the amount of change that may have occurred on the *lissoirs* over 50,000 years, while developing a hypothesis for how they changed. We realize that a substantial amount of methodological development including experimental analyses on possible post-depositional alterations is still needed to accurately reconstruct the surface texture of ancient artifacts. Because it is likely that the two ancient *specimen types* were affected by similar taphonomic processes, it should be reasonable to assume that their surface textures were altered in relatively similar ways. Having the ability to measure multiple surface features allows us to quantify the changes that have occurred over this time frame and may be useful in predicting the initial state of an ancient worked bone. Therefore, the amount of differences between the modern and ancient unworked rib samples should be relatively comparable to the differences between the ancient *lissoir* specimens and what their initial state would have been 50 ka. Using the pairwise comparison for *Sa* and *Spc* as an example, we predict that both of these values have increased over time since the initial deposition of the *lissoirs* (Fig. 5; Fig. 6). Because the ancient unworked ribs tend to be rougher with sharper surface peaks compared with the modern unworked bones, it is likely that the ancient *lissoirs* observed today are rougher with more pointed surface peaks than they were 50 ka. Therefore, the *lissoirs* that were deposited in their respective archaeological sites were most likely less rough with more rounded surface peaks, values that would have likely been lower than those of the modern unworked ribs (Fig. 6).

The overall results of this study indicate that there is a "*lissoir* effect" that is quantitatively different from the ancient unworked rib fragments of the same archaeological deposits. In addition, the ancient and modern unworked ribs show distinguishing quantifiable features that likely result from taphonomic alterations, a key finding that should be helpful for reconstructing the original state of ancient worked bone. Once an archaeological bone artifact's surface texture is confidently reconstructed, these parameters can then be compared with those of modern experimentally worked bone in an effort to understand how the ancient bone object was manufactured and used. This study provides a first step towards predicting the original surface texture values of the ancient *lissoirs*. With more research into how bone surfaces are altered over time, we should be able to make even more accurate predictions about an artifact's original surface state. Extensive analyses are needed to assess various taphonomic effects. This could include chemical analyses with an effort to correlate diagenetic factors such as collagen loss, increased crystallinity, or dissolution with changes to the microtopography of bone (Hedges 2002; Nielsen-Marsh and Hedges 2000). In addition, experimental analyses assessing how differentially worn surfaces are altered through simulated taphonomic effects will be important for making better predictions about the original

state of ancient bone (Gümrükçü and Pante 2018). Further, extensive experimentation on how varying materials affect bone from differing skeletal elements and species will be key in our ability to quantitatively estimate the use of bone artifacts (Martisius et al. 2018; Watson and Gleason 2016). Continued methodological refinement including selection of multiple meaningful surface texture parameters and combining data from differing sources will increase the accuracy of model predictions about the original state of bone artifacts and how they were used in the past. Even so, the assessment of an overall effect of taphonomy on Middle Paleolithic bone tools provides a starting point for future analyses of ambiguous bone artifacts.

Conclusions

Quantitative comparisons between the surfaces of purported archaeological bone tools and experimentally manufactured and used bones are complicated by taphonomic processes. This study demonstrates how unworked bones can be used to quantify the taphonomic effect on bone surfaces and how this effect can then be controlled for and incorporated into an analysis for evaluating the modified surfaces of purported bone tools. Using insights from a previous experimental study that employed a similar Bayesian statistical model and surface texture parameters to bone surfaces (Martisius et al. 2018), we are able to demonstrate the quantitative differences of three *specimen types* (modern unworked ribs, ancient unworked ribs, and ancient *lissoirs*) and use this information to predict surface texture values for the initial state of Middle Paleolithic bone tools typed as *lissoirs*. Importantly, our results demonstrate that modern and ancient unworked ribs are distinguishable and that taphonomic alterations are quantifiable using three of five tested surface texture parameters (Sa, Spc, and IsT). These differences have implications for directly applying surface texture values from experimentally worked bones to archaeological specimens, which could lead to erroneous results. Modified bones from archaeological contexts often have complex sequences of alterations that are further complicated by overlying taphonomic effects. With the ability to quantify taphonomic alterations after deposition by assessing the differences between modern and ancient bones of the same skeletal element, we develop a framework for estimating initial bone artifact surface textures. Crucially, this outcome brings us one step closer to meaningfully applying surface texture data of modern experimental samples to archaeological specimens in order to evaluate how ancient bone objects were manufactured and used. In addition, our results show that there are quantifiable differences between the surfaces of the *lissoirs* and ancient unworked ribs, which corroborate qualitative observations. Results indicate that *lissoir* surfaces are less rough, have more rounded surface peaks, and are more directionally oriented

compared with the unworked rib fragments in the archaeological deposits. This same trend is observed for the predicted initial *lissoir* surface states and the modern ribs, though more variance exists due to uncertainty in the combined model predictions. These differences indicate that the *lissoirs* were anthropogenically modified through extended use, which may include secondary factors like handing. Though we did not set out to infer the function of the Middle Paleolithic artifacts in this study, our results could be consistent with a previous investigation that concluded the use on a soft, supple material for one of the artifacts (Soressi et al. 2013). As a proof of concept, this study provides an important methodological starting point for applying 3D surface texture analysis to archaeological bone artifacts. This analysis of differentially preserved Middle Paleolithic bones compared with modern bone surfaces is a first step towards assessing more ambiguous bone artifacts. Continued methodological refinement will contribute to our ability to accurately interpret ancient bone artifacts and ultimately contribute to our understanding of ancient human technological choices and decision-making.

Data availability The dataset is available within the supplementary material.

Compliance with ethical standards

Conflict of interest The authors have declared that no competing interests exist.

Code availability R script and stan model files are available within the supplementary material.

References

Amemiya Y (1994) On multivariate mixed model analysis. Multivar Anal Appl 24:83–95. https://doi.org/10.1214/lnms/1215463787

Archer W, Braun DR (2013) Investigating the signature of aquatic resource use within Pleistocene hominin dietary adaptations. PLoS One 8:e69899. https://doi.org/10.1371/journal.pone.0069899

Averbouh A (2000) Technologie de la matière osseuse travaillée et implications paléthnologiques : l'exemple des chaînes d'exploitation du bois de cervidé chez les Magdaléniens des Pyrénées. PhD, Université de Paris I-Panthéon Sorbonne

Backwell L, d'Errico F (2001) Evidence of termite foraging by Swartkrans early hominids. Proc Natl Acad Sci U S A 98:1358–1363. https://doi.org/10.1073/pnas.021551598

Baruffi F, Parenti P, Cacciatore F, Annoni M, Tosello G (2017) On the application of replica molding technology for the indirect measurement of surface and geometry of micromilled components. Micromachines 8:195. https://doi.org/10.3390/mi8060195

Behrensmeyer AK (1978) Taphonomic and ecologic information from bone weathering. Paleobiology 4:150–162

Bello SM, Soligo C (2008) A new method for the quantitative analysis of cutmark micromorphology. J Archaeol Sci 35:1542–1552. https://doi.org/10.1016/j.jas.2007.10.018

Bello SM, Verveniotou E, Cornish L, Parfitt SA (2011) 3-dimensional microscope analysis of bone and tooth surface modifications: comparisons of fossil specimens and replicas. Scanning 33:316–324

Beyries S, Delamare F and Quantin J-C (1988) Tracéologie et rugosimétrie tridimensionnelle. Industries lithiques: tracéologie et technologie. British Archaeological Reports Oxford, pp. 115-132.

Böhm K, Winkler DE, Kaiser TM, Tütken T (2019) Post-mortem alteration of diet-related enamel surface textures through artificial biostratinomy: a tumbling experiment using mammal teeth. Palaeogeogr Palaeoclimatol Palaeoecol 518:215–231

Bradfield J (2020) The perception of gloss: a comparison of three methods for studying intentionally polished bone tools. J Archaeol Sci Rep 32:102425

Bradfield J, Forssman T, Spindler L, Antonites AR (2018) Identifying the animal species used to manufacture bone arrowheads in South Africa. Archaeol Anthropol Sci 11:1–16

Burroni D, Donahue RE, Pollard AM, Mussi M (2002) The surface alteration features of flint artefacts as a record of environmental processes. J Archaeol Sci 29:1277–1287

Calandra I, Pedergnana A, Gneisinger W, Marreiros J (2019a) Why should traceology learn from dental microwear, and vice-versa? J Archaeol Sci 110:105012

Calandra I, Schunk L, Bob K, Gneisinger W, Pedergnana A, Paixao E, Hildebrandt A, Marreiros J (2019b) The effect of numerical aperture on quantitative use-wear studies and its implication on reproducibility. Sci Rep 9:6313

Camps-Fabrer H (1966) Matière et Art Mobilier dans la Préhistoire Nord-Africaine et Saharienne. AMPG, Paris

Caux S, Galland A, Queffelec A, Bordes JG (2018) Aspects and characterization of chert alteration in an archaeological context: a qualitative to quantitative pilot study. J Archaeol Sci Reps 20:210–219. https://doi.org/10.1016/j.jasrep.2018.04.027

Currey JD (2006) Bones: structure and mechanics. Princeton university press, Princeton

Curwen EC (1930) Prehistoric flint sickles. Antiquity 4:179–186

d'Errico F, Backwell L (2009) Assessing the function of early hominin bone tools. J Archaeol Sci 36:1764–1773. https://doi.org/10.1016/j.jas.2009.04.005

d'Errico F, Backwell LR, Wadley L (2012) Identifying regional variability in Middle Stone Age bone technology: the case of Sibudu Cave. J Archaeol Sci 39:2479–2495. https://doi.org/10.1016/j.jas.2012.01.040

Dominguez-Rodrigo M, Pickering TR, Bunn HT (2010) Configurational approach to identifying the earliest hominin butchers. Proc Natl Acad Sci U S A 107:20929–20934. https://doi.org/10.1073/pnas.1013711107

Evans AA, Donahue RE (2008) Laser scanning confocal microscopy: a potential technique for the study of lithic microwear. J Archaeol Sci 35:2223–2230. https://doi.org/10.1016/j.jas.2008.02.006

Evans AA, Macdonald D (2011) Using metrology in early prehistoric stone tool research: further work and a brief instrument comparison. Scanning 33:294–303. https://doi.org/10.1002/sca.20272

Galland A, Queffelec A, Caux S, Bordes J-G (2019) Quantifying lithic surface alterations using confocal microscopy and its relevance for exploring the Châtelperronian at La Roche-à-Pierrot (Saint-Césaire, France). J Archaeol Sci 104:45–55

Gaudzinski S (1999) Middle Palaeolithic bone tools from the open-air site Salzgitter-Lebenstedt (Germany). J Archaeol Sci 26:125–141. https://doi.org/10.1006/jasc.1998.0311

Gifford-Gonzalez D (2018) An introduction to zooarchaeology. Springer, Berlin

Giusca C, Evans A, Macdonald D, Leach R (2012) The effect of use duration on surface roughness measurements of stone tools. NPL Report ENG. National Physical Laboratories, Teddington

Goodall RH, Darras LP, Purnell MA (2015) Accuracy and precision of silicon based impression media for quantitative areal texture analysis. Sci Rep 5:10800. https://doi.org/10.1038/srep10800

Grace R (1989) Interpreting the function of stone tools: the quantification and computerisation of microwear analysis.

Grace R, Graham IDG, Newcomer MH (1985) The quantification of microwear polishes. World Archaeol 17:112–120. https://doi.org/10.1080/00438243.1985.9979954

Gümrükçü M, Pante MC (2018) Assessing the effects of fluvial abrasion on bone surface modifications using high-resolution 3-D scanning. J Archaeol Sci Rep 21:208–221

Harris JA, Marean CW, Ogle K, Thompson J (2017) The trajectory of bone surface modification studies in paleoanthropology and a new Bayesian solution to the identification controversy. J Hum Evol 110:69–81. https://doi.org/10.1016/j.jhevol.2017.06.011

Hedges RE (2002) Bone diagenesis: an overview of processes. Archaeometry 44:319–328. https://doi.org/10.1111/1475-4754.00064

Henshilwood CS, d'Errico F, van Niekerk KL, Dayet L, Queffelec A, Pollarolo L (2018) An abstract drawing from the 73,000-year-old levels at Blombos Cave, South Africa. Nature 562:115. https://doi.org/10.1038/s41586-018-0514-3

Hublin JJ, Sirakov N, Aldeias V, Bailey S, Bard E, Delvigne V, Endarova E, Fagault Y, Fewlass H, Hajdinjak M, Kromer B, Krumov I, Marreiros J, Martisius NL, Paskulin L, Sinet-Mathiot V, Meyer M, Paabo S, Popov V, Rezek Z, Sirakova S, Skinner MM, Smith GM, Spasov R, Talamo S, Tuna T, Wacker L, Welker F, Wilcke A, Zahariev N, McPherron SP, Tsanova T (2020) Initial Upper Palaeolithic Homo sapiens from Bacho Kiro Cave, Bulgaria. Nature 581:299–302. https://doi.org/10.1038/s41586-020-2259-z

Ibáñez J, González-Urquijo J, Gibaja J (2014) Discriminating wild vs domestic cereal harvesting micropolish through laser confocal microscopy. J Archaeol Sci 48:96–103

Ibáñez JJ, Anderson PC, Gonzalez-Urquijo J, Gibaja J (2016) Cereal cultivation and domestication as shown by microtexture analysis of sickle gloss through confocal microscopy. J Archaeol Sci 73:62–81

Ibáñez J, Lazuen T, González-Urquijo J (2018) Identifying experimental tool use through confocal microscopy. J Archaeol Method Theory 26:1–40

International Organization for Standardization (2012) ISO 25178-2—geometrical product specifications (GPS)—surface texture: areal—part 2: terms, definitions and surface texture parameters.

Julien MA, Hardy B, Stahlschmidt MC, Urban B, Serangeli J, Conard NJ (2015) Characterizing the Lower Paleolithic bone industry from Schoningen 12 II: a multi-proxy study. J Hum Evol 89:264–286. https://doi.org/10.1016/j.jhevol.2015.10.006

Kaiser TM, Clauss M, Schulz-Kornas E (2016) A set of hypotheses on tribology of mammalian herbivore teeth. Surface Topography: Metrology and Properties 4:014003. https://doi.org/10.1088/2051-672X/4/1/014003

Karr LP, Outram AK (2015) Bone degradation and environment: understanding, assessing and conducting archaeological experiments using modern animal bones. Int J Osteoarchaeol 25:201–212. https://doi.org/10.1002/oa.2275

Krajcarz MT (2019) Alteration of the metal content in animal bones after 2.5-year experimental exposure to sediments. Archaeol Anthropol Sci 11:361–372. https://doi.org/10.1007/s12520-017-0533-2

Langley MC, O'Connor S, Aplin K (2016) A > 46,000-year-old kangaroo bone implement from Carpenter's Gap 1 (Kimberley, northwest Australia). Quat Sci Rev 154:199–213. https://doi.org/10.1016/j.quascirev.2016.11.006

Lartet E (1861) Nouvelles recherches sur la coexistence de l'Homme et des grands Mammifères fossiles réputés caractéristiques de la dernière période géologique. Ann Sci Nat 15:177–253

Legrand A (2007) Fabrication et utilisation de l'outillage en matieres osseuses du Neolithique de Chypre : Khirokitia et Cap Andreas-Kastros. BAR International Series.

Leroi-Gourhan A (1968) Dictionnaire de la Préhistoire. Presses Universitaires de France, Paris

Lesnik JJ (2011) Bone tool texture analysis and the role of termites in the diet of South African hominids. PaleoAnthropology 268-281

Lyman RL (2004) The concept of equifinality in taphonomy. Journal of Taphonomy 2:15–26

Macdonald DA, Harman R, Evans AA (2018) Replicating surface texture: preliminary testing of molding compound accuracy for surface measurements. J Archaeol Sci Rep 18:839–846

Macdonald DA, Xie L, Gallo T (2019) Here's the dirt: first applications of confocal microscopy for quantifying microwear on experimental ground stone earth working tools. J Archaeol Sci Rep 26:101861

Mania D and Mania U (2005) The natural and socio-cultural environment of Homo erectus at Bilzingsleben, Germany. The hominid individual in context: archaeological investigations of Lower and Middle Palaeolithic landscapes, locales and artefacts 98-114.

Martisius NL (2019) The Middle Paleolithic lissoirs (smoothers) from Abri Peyrony and Pech-de-l'Azé I: investigating Neandertal bone tool material selection, manufacture, and use. PhD, University of California, Davis

Martisius NL, Sidéra I, Grote MN, Steele TE, McPherron SP, Schulz-Kornas E (2018) Time wears on: assessing how bone wears using 3D surface texture analysis. PLoS One 13:e0206078. https://doi.org/10.1371/journal.pone.0206078

Martisius NL, Welker F, Dogandzic T, Grote MN, Rendu W, Sinet-Mathiot V, Wilcke A, McPherron SJP, Soressi M, Steele TE (2020) Non-destructive ZooMS identification reveals strategic bone tool raw material selection by Neandertals. Sci Rep 10:7746. https://doi.org/10.1038/s41598-020-64358-w

McPherron SP, Alemseged Z, Marean CW, Wynn JG, Reed D, Geraads D, Bobe R, Bearat HA (2010) Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. Nature 466:857–860. https://doi.org/10.1038/nature09248

Mihlbachler MC, Foy M, Beatty BL (2019) Surface replication, fidelity and data loss in traditional dental microwear and dental microwear texture analysis. Sci Rep 9:1595. https://doi.org/10.1038/s41598-018-37682-5

Mons L, Stordeur D (1977) Des objets nommés "lissoirs" de la Grotte du Placard (Charente). Antiquités Nationales St-Germain-en-Laye 9: 15–25

Moore KM (2013) Economic and social context of bone tool use, formative Bolivia. In: Choyke A, O'Connor S (eds) From these bare bones: raw materials and the study of worked osseous objects. Oxbow Books, Oxford, pp 174–187

Murdoch D, Chow ED and Celayeta JMF (2013) Functions for drawing ellipses and ellipse-like confidence regions.

Nielsen-Marsh CM, Hedges RE (2000) Patterns of diagenesis in bone I: the effects of site environments. J Archaeol Sci 27:1139–1150. https://doi.org/10.1006/jasc.1999.0537

Orłowska J (2018) The same or different? Experimenting with the influence of peat environment on use-wear traces on antler tools. Int J Osteoarchaeol 28:120–130

Otárola-Castillo E, Torquato MG (2018) Bayesian statistics in archaeology. Annu Rev Anthropol 47:435–453

Otárola-Castillo E, Torquato MG, Hawkins HC, James E, Harris JA, Marean CW, McPherron SP, Thompson JC (2018) Differentiating between cutting actions on bone using 3D geometric morphometrics and Bayesian analyses with implications to human evolution. J Archaeol Sci 89:56–67

Pante MC, Muttart MV, Keevil TL, Blumenschine RJ, Njau JK, Merritt SR (2017) A new high-resolution 3-D quantitative method for identifying bone surface modifications with implications for the Early Stone Age archaeological record. J Hum Evol 102:1–11. https://doi.org/10.1016/j.jhevol.2016.10.002

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Radmilli A and Boschian G (1996) Gli scavi a Castel di Guido. Il più antico giacimento di cacciatori nell'Agro Romano. ORIGINES

Rho JY, Kuhn-Spearing L, Zioupos P (1998) Mechanical properties and the hierarchical structure of bone. Med Eng Phys 20:92–102

Rosso DE, d'Errico F, Queffelec A (2017) Patterns of change and continuity in ochre use during the late Middle Stone Age of the Horn of Africa: the Porc-Epic Cave record. PLoS One 12:e0177298. https://doi.org/10.1371/journal.pone.0177298

Roth M (2013) On the multivariate t distribution. Linköping University Electronic Press, Linköping

Sano K, Beyene Y, Katoh S, Koyabu D, Endo H, Sasaki T, Asfaw B, Suwa G (2020) A 1.4-million-year-old bone handaxe from Konso, Ethiopia, shows advanced tool technology in the early Acheulean. Proc Natl Acad Sci U S A 117:202006370. https://doi.org/10.1073/pnas.2006370117

Schulz E, Calandra I, Kaiser TM (2010) Applying tribology to teeth of hoofed mammals. Scanning 32:162–182. https://doi.org/10.1002/sca.20181

Schulz E, Calandra I, Kaiser TM (2013a) Feeding ecology and chewing mechanics in hoofed mammals: 3D tribology of enamel wear. Wear 300:169–179. https://doi.org/10.1016/j.wear.2013.01.115

Schulz E, Piotrowski V, Clauss M, Mau M, Merceron G, Kaiser TM (2013b) Dietary abrasiveness is associated with variability of microwear and dental surface texture in rabbits. PLoS One 8: e56167. https://doi.org/10.1371/journal.pone.0056167

Semenov SA (1964) Prehistoric technology. Cory, Adams & Mackay, London

Sidéra I (1993) Les assemblages osseux en bassins parisien et rhénan du VIe au IVe millénaire B.C. Histoire, techno-économie et culture. PhD, Université de Paris I

Sonneville-Bordes Dd (1960) Le Paléolithique Supérieur en Périgord. Bordeaux.

Soressi M, McPherron SP, Lenoir M, Dogandzic T, Goldberg P, Jacobs Z, Maigrot Y, Martisius NL, Miller CE, Rendu W, Richards M, Skinner MM, Steele TE, Talamo S, Texier JP (2013) Neandertals made the first specialized bone tools in Europe. Proc Natl Acad Sci U S A 110:14186–14190. https://doi.org/10.1073/pnas.1302730110

Stan Development Team (2019) RStan: the R interface to Stan. R package version 2.19.1 ed.

Stemp WJ, Chung S (2011) Discrimination of surface wear on obsidian tools using LSCM and RelA: pilot study results (area-scale analysis of obsidian tool surfaces). Scanning 33:279–293. https://doi.org/10.1002/sca.20250

Stemp WJ, Stemp M (2001) UBM laser profilometry and lithic use-wear analysis: a variable length scale investigation of surface topography. J Archaeol Sci 28:81–88. https://doi.org/10.1006/jasc.2000.0547

Stemp WJ, Stemp M (2003) Documenting stages of polish development on experimental stone tools: surface characterization by fractal geometry using UBM laser profilometry. J Archaeol Sci 30:287–296. https://doi.org/10.1006/jasc.2002.0837

Stemp WJ, Morozov M, Key AJM (2015) Quantifying lithic microwear with load variation on experimental basalt flakes using LSCM and area-scale fractal complexity (Asfc). Surf Topogr 3:034006. https://doi.org/10.1088/2051-672x/3/3/034006

Stemp WJ, Watson AS, Evans AA (2016) Surface analysis of stone and bone tools. Surface Topography: Metrology and Properties 4. https://doi.org/10.1088/2051-672x/4/1/013001

Stevens NE, Harro DR, Hicklin A (2010) Practical quantitative lithic use-wear analysis using multiple classifiers. J Archaeol Sci 37:2671–2678. https://doi.org/10.1016/j.jas.2010.06.004

Tartar É (2009) De l'os à l'outil: Caractérisation technique, économique et sociale de l'utilisation de l'os à l'Aurignacien ancien: Étude de trois sites: l'Abri Castanet (secteurs nord et sud), Brassempouy (Grotte des Hyènes et Abri Dubalen) et Gatzarria. PhD, Université de Paris I-Panthéon-Sorbonne

Tejero J-M (2014) Towards complexity in osseous raw material exploitation by the first anatomically modern humans in Europe: Aurignacian antler working. J Anthropol Archaeol 36:72–92

Tejero J-M, Yeshurun R, Barzilai O, Goder-Goldberger M, Hershkovitz I, Lavi R, Schneller-Pels N, Marder O (2016) The osseous industry from Manot Cave (Western Galilee, Israel): technical and conceptual behaviours of bone and antler exploitation in the Levantine Aurignacian. Quat Int 403:90–106

Valtierra N, Courtenay LA, López-Polín L (2020) Microscopic analyses of the effects of mechanical cleaning interventions on cut marks. Archaeol Anthropol Sci 12:193. https://doi.org/10.1007/s12520-020-01153-8

Vaughan PC (1981) Lithic microwear experimentation and the functional analysis of a Lower Magdalenian stone tool assemblage.

Vehtari A, Gelman A, Jonah G (2017) Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat Comput 27:1413–1432

Vietti LA (2016) Quantifying bone weathering stages using the average roughness parameter Ra measured from 3D data. Surf Topogr Metrol Prop 4:034006. https://doi.org/10.1088/2051-672x/4/3/034006

Watson AS, Gleason MA (2016) A comparative assessment of texture analysis techniques applied to bone tool use-wear. Surf Topogr Metrol Prop 4:024002

Werner JJ (2018) An experimental investigation of the effects of post-depositional damage on current quantitative use-wear methods. J Archaeol Sci Rep 17:597–604. https://doi.org/10.1016/j.jasrep.2017.12.008

Zhang SQ, d'Errico F, Backwell LR, Zhang Y, Chen FY, Gao X (2016) Ma'anshan cave and the origin of bone tool technology in China. J Archaeol Sci 65:57–69. https://doi.org/10.1016/j.jas.2015.11.004

Zhang SQ, Doyon L, Zhang Y, Gao X, Chen FY, Guan Y, d'Errico F (2018) Innovation in bone technology and artefact types in the late Upper Palaeolithic of China: insights from Shuidonggou Locality 12. J Archaeol Sci 93:82–93. https://doi.org/10.1016/j.jas.2018.03.003