Detecting skill level and mental templates in Late Acheulean biface morphology: Archaeological and experimental insights

Cheng Liu* Nada Khreisheh† Radu Iovita‡ Dietrich Stout§ Justin Pargeter¶

Abstract

Despite the extensive literature focusing on Acheulean bifaces, especially the sources and meaning of their morphological variability, many aspects of this topic remain elusive. Among many factors identified to contribute to the considerable variation of biface morphology, skill level and mental templates have been frequently cited. Here we present results from a multidisciplinary study of Late Acheulean handaxe-making skill acquisition involving twenty-six naïve participants trained for up to 90 hours in Late Acheulean style handaxe production and three expert knappers. We compare their handaxe to the Late Acheulean handaxe assemblage from Boxgrove, UK. Through the principal component analysis of morphometric data derived from images, our study suggested that both skill level and mental template have a relatively clear manifestation in different aspects of biface morphology, where the former is related to cross-sectional thinning (PC1) while the latter relates to handaxe elongation and pointedness (PC2). Moreover, we also evaluated the effects of training using the data from a 90-hour long knapping skill acquisition experiment and found that reaching the skill level of modern expert requires more training time than was permitted in this extensive and long-running training program. Our study demonstrated the potential of experimental archaeology and digital photographs in revealing new insights from old archaeological assemblages.

Keywords: Late Acheulean; Biface production; Boxgrove; Experimental archaeology; Skill level; Mental template

Contents

*Department of Anthropology, Emory University, Atlanta, GA, USA; raylc1996@outlook.com

†The Ancient Technology Centre, Cranborne, Dorset, UK; nada.khreisheh@dorsetcouncil.gov.uk

‡Department of Anthropology, New York University, New York, NY, USA; Department of Early Prehistory and Quaternary Ecology, Eberhard Karls University of Tübingen, Tübingen, Germany; iovita@nyu.edu

§Department of Anthropology, Emory University, Atlanta, GA, USA; dwstout@emory.edu

¶Department of Anthropology, New York University, New York, NY, USA; Palaeo-Research Institute, University of Johannesburg, Auckland Park, South Africa; justin.pargeter@nyu.edu

1 Introduction

The morphological variability of Acheulean bifaces has been one of the most well-studied and well-published topics in paleolithic archaeology (Key & Lycett, 2019; Petraglia & Korisettar, 1998; White, 1998). Despite the recurrent narrative emphasizing the heterogeneity and longevity of biface assemblage on a global scale and the conservatism behind this phenomenon that evokes genetic explanations (Corbey et al., 2016; Corbey, 2020; Richerson & Boyd, 2005; Sterelny, 2004), many researchers have recognized the diversity within what has been deemed as a unified Acheulean "tradition" and tried to dissect the sources and meaning of this variation (Lycett & Gowlett, 2008; Nowell, 2002; Nowell & White, 2010; Sharon et al., 2011). More specifically, a complex suite of interconnecting factors have been identified to contribute to the great variation of biface morphology, including but not limited to raw materials (Eren et al., 2014; Lycett et al., 2016; Sharon, 2008), percussor properties (Shipton et al., 2009), function (Key et al., 2016; Key & Lycett, 2017; Kohn & Mithen, 1999; Machin et al., 2007; White & Foulds, 2018), reduction method/intensity (Shipton et al., 2009; Shipton & Clarkson, 2015), learning processes (Kempe et al., 2012; Lycett et al., 2016), skill level (Caruana & Herries, 2021; Herzlinger et al., 2017; Stout et al., 2014), mental template (García-Medrano et al., 2019; Hutchence & Scott, 2021). From this extensive list, skill level and mental template have been repeatedly mentioned and discussed in the now extensive corpus of biface studies.

In its classical definition, mental template is the "idea of the proper form of an object exists in the mind of the maker, and when this idea is expressed in tangible form in raw material, an artifact results" (Deetz, 1967: 45). This concept essentially lies in the very foundation of the cultural-historical approach in that the identification of archaeological cultures is based on the existence of distinct mental templates in a given spatial-temporal framework. Early researchers, whether explicitly or implicitly, often endorse this conceptual framework and actively applies in the typological analysis of bifaces at the regional level (Roe, 1969; Wenban-Smith et al., 2000; Wenban-Smith, 2004). Combined with the production of large flakes, the emergence of mental templates (or "imposed form") has been recognized as two major technological innovations of the Acheulean compared with the Oldowan (Isaac, 1986). For a decade or so, this concept has been less frequently used, since it was criticized for its normative and static assumption (Lyman & O'Brien, 2004), ignorance of other competing factors such as raw material constraints (White, 1995), and the lack of rigorous studies of its corresponding cognitive processes. Until very recently, several researchers have actively addressed these critiques and reconceptualized the mental template in the study of biface morphology. Hutchence and Scott (2021: 675), for example, leveraged the theory of "community of practice" (Wenger, 1998) and argued that "the consistency in handaxe shape found at sites like Boxgrove is a consequence of the emergent actions of individual knappers being simultaneously constrained by the imposition of social norms." Furthermore, raw material is often treated as a crucial variable to be controlled at the very beginning of a research design focusing on mental template. This is best exemplified by an experimental study of García-Medrano et al. (2019), where they carefully chose experimental nodules mirroring those found in archaeological context in composition, size, and shape. In terms of the cognitive mechanisms behind mental template, Ho and colleagues (2022) recently developed a series of navigation experiments demonstrating the externalization of the planning process to simple geometric representations, which has the potential to be transferred into a research setting aiming at directly testing the planning of knapping behaviors. In short, when exercised with proper caution, mental template still has its value in our study of biface morphological variation, which can be further dissected into a series of shape variables corresponding to pointedness and elongation, among other things.

Building upon the concept of mental template, one possible way of defining skill is the capacity for a knapper to realize mental templates using the resources available (Roux et al., 1995: 66). This version of conceptualization, particularly relevant when it comes to motor skills such as knapping, can be dismantled into two mutually dependent aspects, namely the intentional aspect (goal/strategic planning) and the operational aspect (means/motor execution) (Connolly

& Dalgleish, 1989). It also roughly corresponds to the well-known dichotomy developed by French lithic analysts of "*connaissance*" (abstract knowledge) and "*savoir-faire*" (practical know-how) (Pelegrin, 1993). As Stout (2002: 694) noted, the acquisition of skill is deeply rooted in its social context, and it is not composed of "some rigid motor formula" but "how to act in order to solve a problem." This ecological notion of skill somewhat mirrors Hutchence and Scott's (2021) reconceptualization of the mental template in that they both refute the idea that technology is simply an internal program expressed by the mind and they prefer a dynamic approach emphasizing the interaction between perception and action. The manifestations of skill in materialized form display a great amount of variation, but ethnoarchaeological studies have repeatedly suggested that skills can be improved through practice as perceived by the local practitioners. It is thus possible to evaluate the skill levels reflected in knapping products (Roux et al., 1995; Stout, 2002). When contextual information is less readily available as in the Late Acheulean archaeological assemblages, how to properly operationalize and measure knapping skills has been a methodological issue receiving much attention among archaeologists (Bamforth & Finlay, 2008; Kolhatkar, 2022). In addition to measurements that can be almost applied in any lithic technological system such as raw materials, platform preparation, as well as hinges, in the context of biface technology, symmetry (Hodgson, 2015; Hutchence & Debackere, 2019) and cross-sectional thinning (Caruana, 2020; Pargeter et al., 2019; Stout et al., 2014) have been frequently quoted as reliable and distinctive indicators of the skill level as supported by several experimental studies. These two features have also been commonly used as standards for dividing Early Acheulean and Late Acheulean (Callahan, 1979; Clark, 2001; Schick & Toth, 1993).

Drawing on these two lines of literature, we aim to explore the possibility of differentiating skill level and mental template through a comparative study of an archaeological biface assemblage known for its remarkable dexterity, a reference biface collection produced by modern knapping experts, and an experimental biface sample produced by modern novice knappers. Since the novice biface collection is generated from a 90-hour skill acquisition experiment, we also have the precious opportunity to introduce the diachronic dimension of training time and interrogate its impact on the variables of interest. As such, we propose the following two interconnected research questions in this article: 1) Can skill level and mental template be efficiently detected from biface morphometric data? 2) How does training affect novices' performance in these two aspects?

2 Materials and methods

2.1 Boxgrove biface collection

The archaeological site of Boxgrove is located in the former Eartham quarry, Boxgrove, West Sussex, featuring a long sequence of Middle Pleistocene deposit (Pope et al., 2020; Roberts & Parfitt, 1998). This 500-ka-old site has documented exceedingly rich details of Lower Paleolithic hominins' subsistence behaviors (Smith, 2013, 2012) and their paleoenvironmental contexts (Holmes et al., 2010). In addition to the presence of one of the earliest hominin fossil (*Homo heidelbergensis*, Hillson et al., 2010) and bone assemblages with anthropogenic modifications in northern Europe (Bello et al., 2009), Boxgrove is mostly known for its large sample size of Late Acheulean-style flint handaxes and the high dexerity reflected in their manufacture. As such, it has received wide research attention in the past two decades regarding the relationships between technology, cognition, and skills (García-Medrano et al., 2019; Iovita et al., 2017; Iovita & McPherron, 2011; Shipton & Clarkson, 2015; Stout et al., 2014). To identify the morphological manifestation of knappers' dexterity in our study, we selected a complete handaxe assemblage (n=326) previously analyzed and reported in digital formats by Iovita and McPherron (Iovita & McPherron, 2011), which is currently curated at the Franks House of the British Museum (Iovita et al., 2017). The digital photographs are taken of each handaxe at a 90° angle, which was oriented with the tip to the right of the photos, and the camera faces the most convex surface of the handaxe (Iovita & McPherron, 2011).

2.2 Experimental biface collection

The biface experimental replicas used in this study comprised two sub-collection. The first sub-collection includes 10 bifaces knapped by three expert knappers, including Bruce Bradley (n=4), John Lord (n=3), and Dietrich Stout (n=3) (Stout et al., 2014). These handaxes were made for previous research projects, which similarly aimed to approximate 'Late Acheulean' handaxes explicitly comparable to the Boxgrove assemblage (Faisal et al., 2010; Stout et al., 2014; Stout et al., 2011). The second sub-collection is produced from a 90-hour handaxe knapping skill acquisition experiment (Bayani et al., 2021; Pargeter et al., 2020; Pargeter et al., 2019), where 30 adults with no previous experience in knapping were recruited from Emory University and its surrounding communities and requested to make 132 bifaces in total. Among these 30 adult participants, 17 have gone through multiple one-to-one or group training sessions that amounted to 89 hours in maximum, while the remaining 13 were assigned to the controlled group, where no formal training is given. As part of the preparation efforts, the experimental team spalled the Norfolk flints acquired through Neolithics.com into flat blanks of similar size and shape for training and assessments. The mechanical properties of these raw materials are comparable to the ones used in Boxgrove in that they are both fine-grained and highly predictable in fracturing process.

In this experiment, all research participants participated in the initial assessment (assessment 1 in our data set) before formal training, where they each produced a handaxe after watching three 15-minute videos of Late Acheulean style handaxes demonstrated by expert knappers and examining four Late Acheulean style handaxe replicas. Training was provided by verbal instruction and support from the second author, an experienced knapping instructor (Khreisheh et al., 2013) with 10 years knapping practice and specific knowledge of Late Acheulean technology including the Boxgrove handaxe assemblage. She was present at all training sessions to provide help and instruction to participants. All training occurred under controlled conditions at the outdoor knapping area of Emory's Paleolithic Technology Lab, with knapping tools and raw materials provided. All participants were instructed in basic knapping techniques including how to select appropriate percussors, initiate flaking on a nodule, maintain the correct flaking gestures and angles, prepare flake platforms, visualize outcomes, deal with raw material imperfections, and correct mistakes. Handaxe-specific instruction included establishment and maintenance of a bifacial plane, cross-sectional thinning, and overall shaping. The training emphasized both aspects of handaxe making technical skill (the importance of producing thin pieces with centered edges) as well as mental template related markers (symmetrical edges).

Subsequently, the 17 participants in the experimental group were assessed after every ten hours of the cumulative learning period, where each of them was requested to produce a handaxe for expert knapper's (N. Khreisheh) review, leading to the compilation of a data set composing 9 assessments in total. It should be also noted that 6 out of 17 participants dropped out of the research before the final assessment due to personal reasons. To detect the effect of training on skill level and mental template, we reorganized our assessment classification scheme and combined it into three broader categories, namely pre-training (assessment 1), early training (assessment 2-5), and late training (assessment 6-9), which helps increase the sample size of the measured intervals. A more detailed experimental protocol can be assessed in one of our published papers (Pargeter et al., 2019).

2.3 Lithic analysis

To better understand the morphological variation of Boxgrove biface collection, we adopted a standardized analytical procedure to extract the morphometric information from 752 photos of the studied samples prepared by R. I. (Iovita & McPherron, 2011), which include both the front and lateral views of a given specimen. First, we used Adobe Photoshop to conduct a batch transformation of the samples' pixel scale into a real-world measurement scale based on the fixed photographic setting. This is then followed by the batch conversion of color photographs to a black-and-white binary format. Subsequently, we cropped the silhouettes of bifaces one by one using the Quick Selection Tool in Adobe Photoshop. The metric measurements were conducted in ImageJ (Rueden et al., 2017), where we employed a custom ImageJ script (Pargeter et al., 2019) to measure the maximum length, width, and thickness of a given silhouette. The width and thickness measurements are taken at 10% increments of length starting at the tip of each biface (**Figure** 1), which eventually leads to 19 morphometric variables in total (1 length measurement, 9 width measurements, and 9 thickness measurements). Finally, we calculated the geometric means of all 19 linear measurements to create a scale-free data set that preserves the individual morphological variation at the same time (Lycett et al., 2006), which also at least alleviate the effect of resharpening process to some extent. The same procedure was also applied to the morphometric analyses of the experimental biface collection, which was partially published in Pargeter et al. (2019).

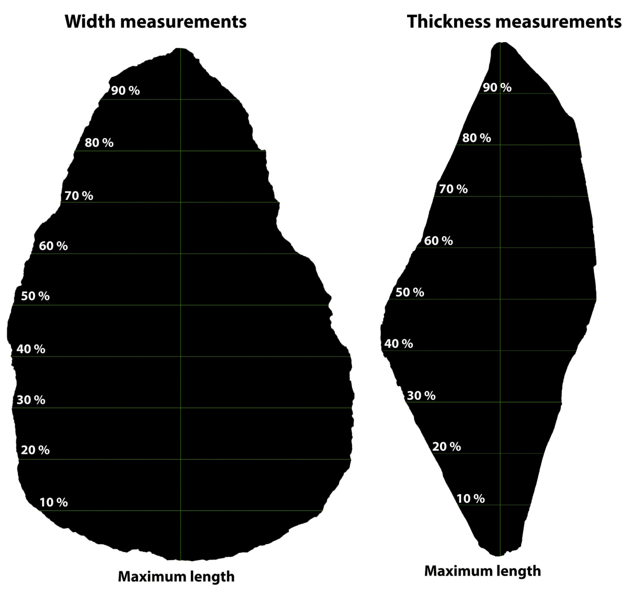

Figure 1: A visual demonstration of the biface measurement protocal using Image J (after Pargeter et al. 2019: Figure 5).

2.4 Statistical analyses

Given the number of variables involved in this study, we used principal component analysis (PCA) to reduce the dimension and identify the possible patterns in this morphometric data set, which is one of the most used techniques in similar studies (García-Medrano, Maldonado-Garrido, et al., 2020; García-Medrano, Ashton, et al., 2020; Herzlinger et al., 2017; Iovita & McPherron, 2011; Shipton & Clarkson, 2015; Stout et al., 2014). To detect the effect of training on novices' performance as compared with archaeological samples and biface made by experts, we also compare the corresponding metrics built on PCA across different training periods and across all groups using the Games-Howell nonparametric post-hoc test, which does not rely on the assumptions of equal sample sizes and equal variance. This study adheres to the principles of reproducibility and data transparency of archaeological research by depositing all the codes and data sets involved in an open-access online repository (Marwick, 2017), which can be accessed through the author's Github (https://github.com/Raylc/PaST-pilot).

3 Results

3.1 Principal component analysis

Our analysis suggested that the first two components already explain 77.2% of the variation for the entire morphometric data set composed of 19 variables (**Figure** 2), which is a rather decent explained variance ratio to avoid overfitting. We then decided to focus on and further interpreted the implications of these first two components based on their relationships between variables (**Table** 1). The first principal component (PC1) indicates the overall cross-sectional thickness as it is positively correlated with all thickness measurements while negatively correlated with all other measurements. That being said, a higher PC1 value indicates a thicker biface, and vice versa. The second principal component (PC2) tracks the elongation and pointedness based on its positive relationship with maximum length and bottom width/thickness. As PC2 increases, a biface will be generally longer and more pointed since its bottom part will be bulkier.

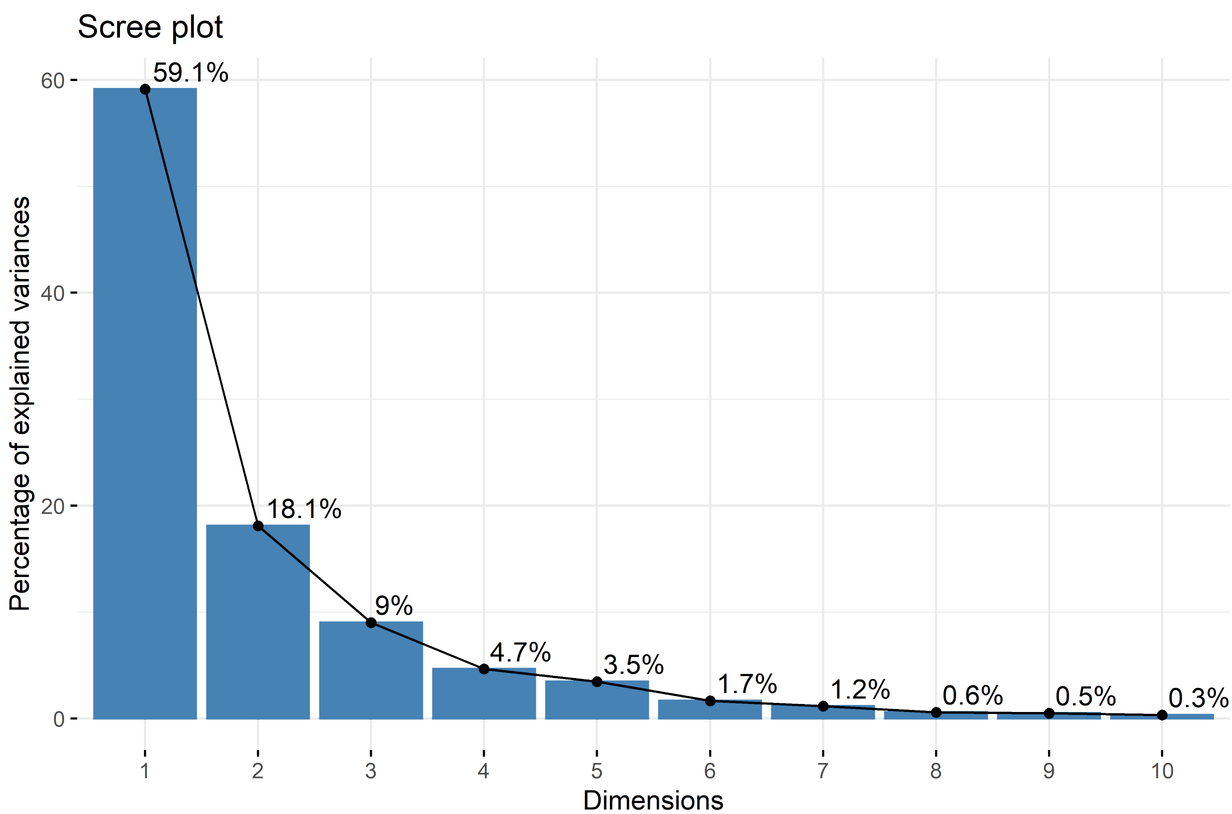

Figure 2: A scree plot showing the percentage of explained variances of the first 10 principal components.

Table 1: Variable loadings for the first two principal components

X	Dim.1	Dim.2
width_0.1	-0.1131312	-0.1256408
width_0.2	-0.1419554	-0.1326946
width_0.3	-0.1684170	-0.1232328
width_0.4	-0.1867226	-0.0966578
width_0.5	-0.2037483	-0.0651505
width_0.6	-0.2121330	-0.0197136
width_0.7	-0.2083163	0.0232790
width_0.8	-0.1885821	0.0661257
width_0.9	-0.1447319	0.0805702
thickness_0.1	0.0142639	-0.0240388
thickness_0.2	0.0247137	-0.0227114
thickness_0.3	0.0435524	-0.0093580
thickness_0.4	0.0667936	0.0047643
thickness_0.5	0.0893523	0.0261202
thickness_0.6	0.1083112	0.0484852
thickness_0.7	0.1288346	0.0628567
thickness_0.8	0.1444047	0.0659257
thickness_0.9	0.1308949	0.0487419
max_length	-0.3626265	0.2507234

A closer look at the principal component scatter plot (**Figure** 3) yields the clustering of different groups of bifaces. The majority of Boxgrove bifaces occupy an area featuring negative values of both PC1 and PC2. The expert group is similar to the Boxgrove group in PC1, while the former has a relatively higher PC2 value than the latter on average. The group of novice displays the highest level of variability, however, it is rather pronounced that most bifaces made by novices have a positive PC1 value that is different from both the groups of Boxgrove and experts.

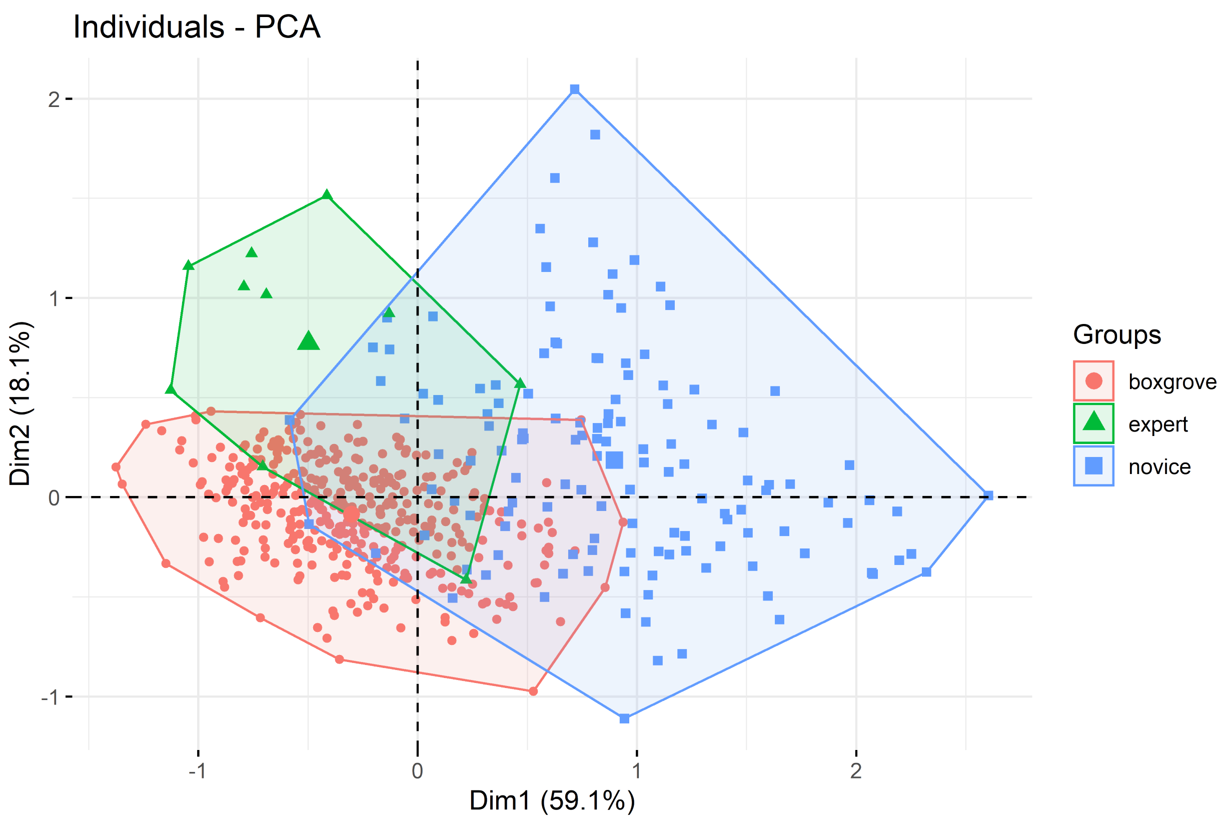

Figure 3: A principal component scatter plot of bifaces from the groups of Boxgrove (red, n=326), expert (green, n=10), and novice (blue, n=132).

3.2 Effects of training

We extracted the PC1 and PC2 values of individual bifaces and compared them between different groups. More specifically, the novice group was divided into three sub-groups based on their training stages as specified in the method section. As such, we found that for PC1 values (**Figure** 4), the only two group comparisons that are **not** statistically significant are the one between Boxgrove and Expert ($t = -1.65$, $p > 0.05$) and the one between Early training and Late training stages ($t = -0.649$, $p > 0.05$), which at least partially confirms our visual observation of the general PCA scatter plot. Likewise, for PC2 values (**Figure** 5), the group comparison between the Early training and Late stages again is **not** statistically significant ($t = 0.333$, $p > 0.05$). However, a rather surprising result here is that the mean PC2 value difference between the Pre-training group and Boxgrove is also **not** statistically significant ($t = -0.818$, $p > 0.05$).

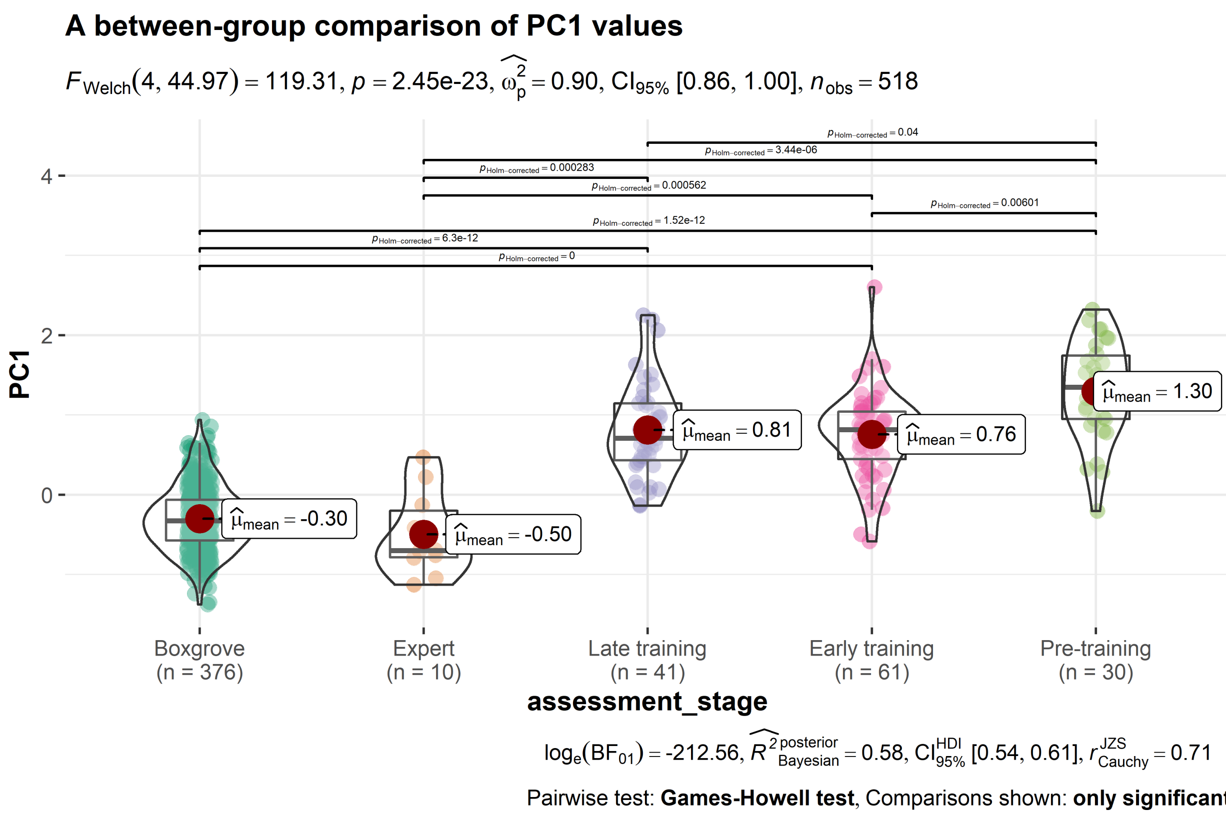

Figure 4: A between-group comparison of PC1 values.

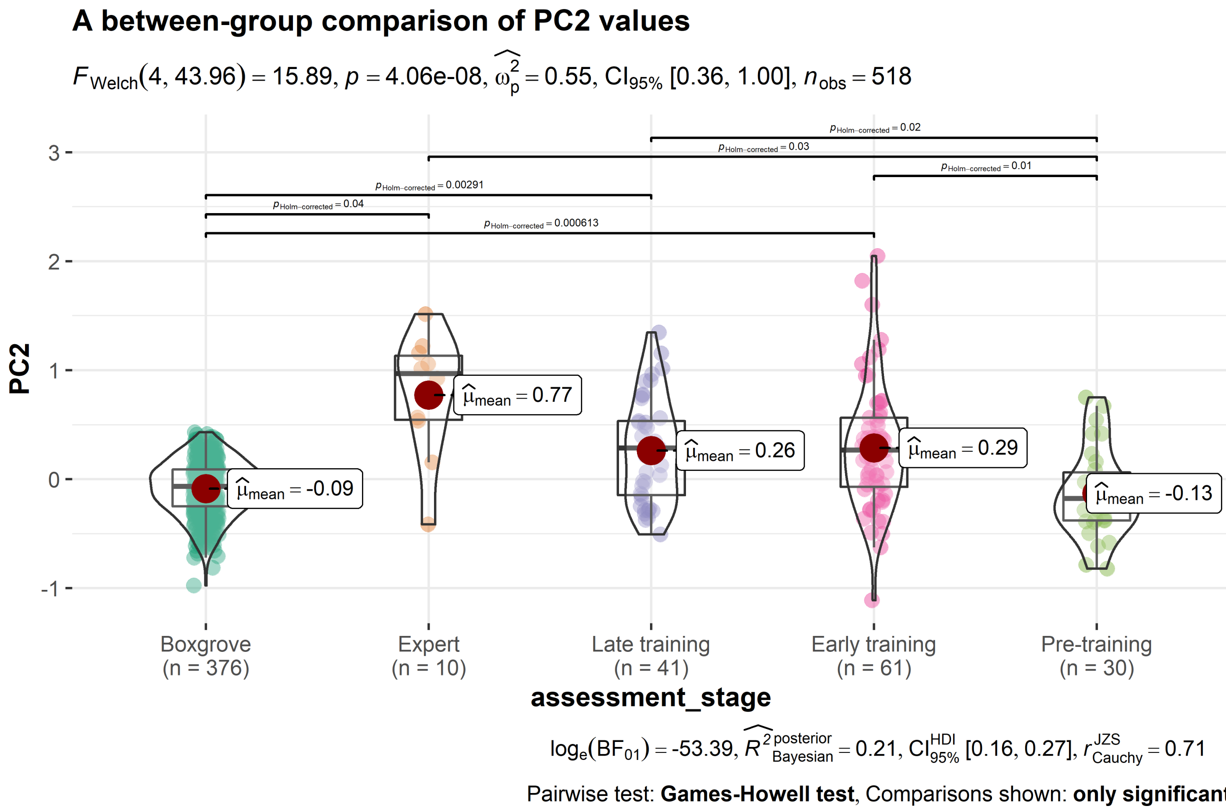

Figure 5: A between-group comparison of PC2 values.

4 Discussion

In accordance with the existing literature on biface knapping skill (Callahan, 1979; Caruana, 2020; Stout et al., 2014), the results of PCA suggested that PC1 (cross-sectional thinning) is a robust indicator of skill level as it is a common feature shared by modern expert knapper and Boxgrove knappers. Thinning is regarded as a technique requiring a high knapping skill level because it requires one to carefully detach flakes in an invasive manner while not breaking the biface into several pieces, serving the purpose of achieving the desired convexity and/or volume. This procedure involves precise control of striking forces, strategic choice of platform external angle, and attentive preparation of bifacial intersection plane, all of which were part of our experimental training program (Callahan, 1979; Caruana, 2022; Pargeter et al., 2020; Shipton et al., 2013; Stout et al., 2014). Experimental studies have also shown that the thinning stage of biface produce often involves the use of soft hammers, which is also supported by indirect archaeological evidence of flake attributes from Boxgrove (Roberts & Parfitt, 1998: 384-394; Roberts & Pope, 2009). This also reflects the majority of samples in both our expert and novice experimental assemblages. In the skill acquisition experiments, novice knappers have been explicitly taught to switch to the soft hammer for thinning purposes, but some of them did not follow the instruction during the assessment. On the other hand, it has also been shown that hard hammers can also be used to achieve similar thinning results (Bradley & Sampson, 1986; Pelcin, 1997), corresponding to the cases of replicas produced by Bruce Bradley and a few novices in our study.

Given the dissimilarity of PC2 (elongation and pointedness) values between archaeological and experimental samples and its similarity among modern knappers, we argue that this dimension reflects different mental templates, where the Boxgrove assemblage displays an ovate shape featuring a wider tip while the experimental assemblages are characterized by a more pointed shape with a longer central axis. This divergence of group-level aesthetic choices can be best explained under the theoretical framework of the communities of practice (Wenger, 1998) as advocated by Hutchence and Scott in biface analysis(2021). The most common form of learning in the experiment occurred in the group condition, where the instructor taught multiple novices at the same time and novices have the chance to also communicated and learned from their peers. Unfortunately, the biface data from the instructor (N. Khreisheh) are unavailable, but it should be noted that the instructor has learned how to knap and how to teach knapping from one of our expert knapper (Bruce Bradley). This cascading effect of social learning might explain why there is a shared mental template between the expert group and the novice group after training.

In terms of our second research question, this study shows that training does have an immediate intervention effect (pre-training vs. post-training) in both PC1 (skill level) and PC2 (mental template). Nonetheless, once the training has been initiated, its effects across different assessments on both dimensions are rather unconsipicous. This finding corroborates what has been suggested in Pargeter et al. (2019) that 90 hours of training for handaxe making is still not enough for novices to reach the skill level as reflected in expert knappers, considering the massive social support involved in the experiment set up including the direct and deliberate pedagogy and the simplified raw material procurement and preparation procedures. This follow-up project further adds the samples produced by the Late Acheulean toolmaker as a new benchmark to deepen our understanding of this issue. It is noteworthy how constrained is the Boxgrove assemblage morphological variation as measured by both PC1 and PC2 even when compared with the modern expert group (**Figure** 3), especially given the fact that it has the largest sample size among all studied group. Some potential explanations for this phenomenon include 1) the strong idiosyncrasy of individual expert knappers shaped by their own unique learning and practice experience; and/or 2) the present day-skill shortage of our expert knapper as compared with Boxgrove knappers despite their multiple years of knapping practice (Milks, 2019).

The pre-training group is similar to the Boxgrove group in PC2 because these novices lack the ability to effectively reduce the nodules, which are typically flat pre-prepared cortical flakes, to the desired form (**Figure** 6). If the given nodules already possess an oval morphology like those presented in the Boxgrove assemblage, it is likely the form of end products knapped by novices in the pre-training group will remain roughly unchanged. This explanation is also supported by the comparison of average delta weight, defined as the difference between the weight of handaxe and the weight of nodule, among four groups, where the pre-training group displays the lowest value (**Figure** 7). On the other hand, the refitting analyses of the Boxgrove handaxe assemblage have suggested that the nodules exploited by knappers inhabiting this site are somewhat bulky and amorphous (Roberts & Parfitt, 1998: 339, 360). These characteristics have been clearly displayed in a recent attempt of slow-motion refitting of a handaxe specimen from Boxgrove GTP17 (https://www.youtube.com/watch?v=iS58MUJ1ZEo). As such, behind the resemblance of the pre-training group and the Boxgrove assemblage in PC2 are two types of mechanisms that are fundamentally different from each other, where the latter group exhibits a complex suite of cognitive and motor execution processes to transform the shapeless raw materials to a delicate end product in a given shape.

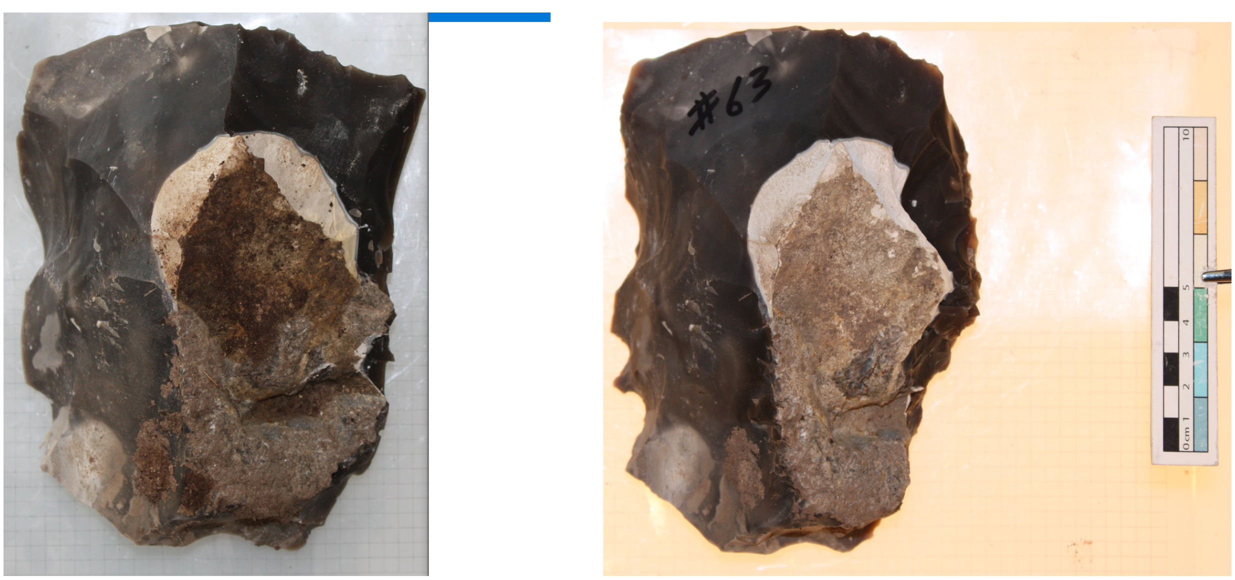

Figure 6: Core 63 before (left) and after knapping(right), showing the minimal morphological change during the process

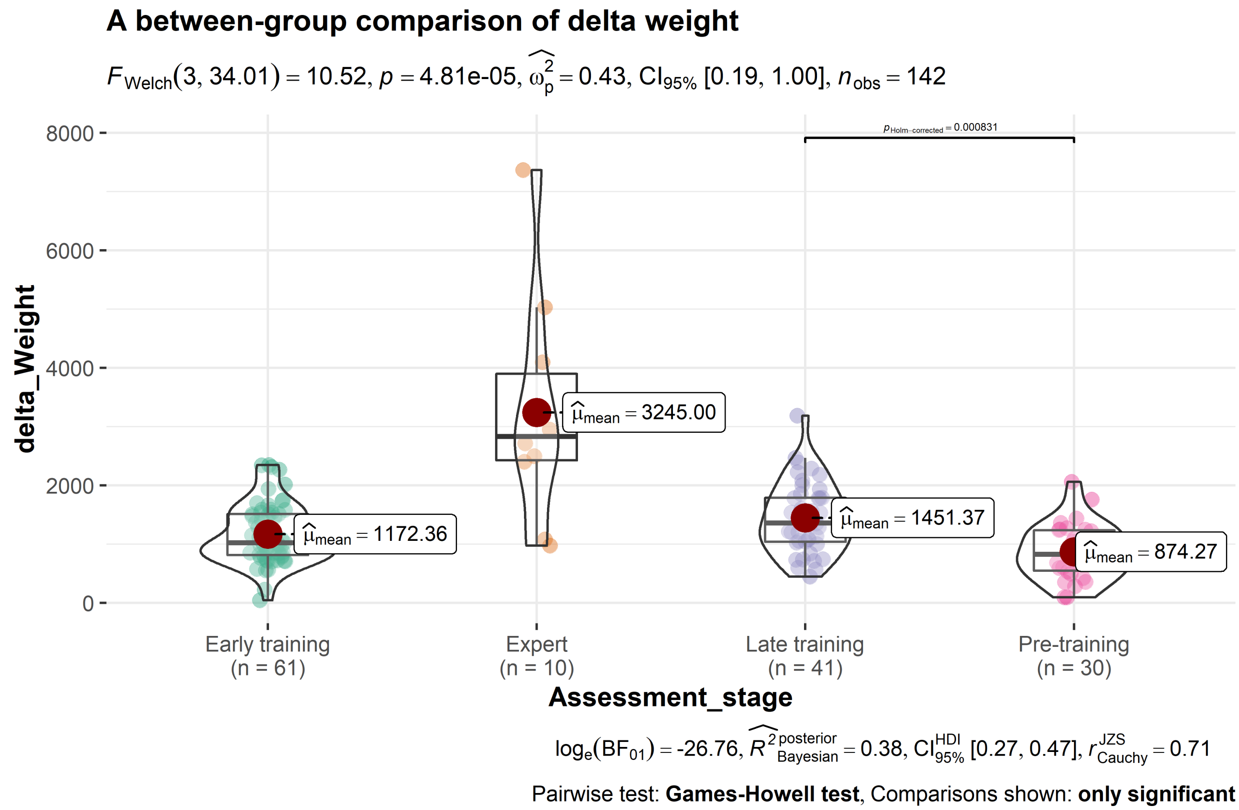

Figure 7: A comparison of the delta weight between the pre-training, early training, late training, and the expert group.

Another contribution that we would like to highlight here is that this research project demonstrates the potential of reusing old archaeological data in digital format to address novel research questions. In this paper, the main source of archaeological data is a collection of photos produced and curated by one of our co-authors (R. Iovita) more than 10 years ago, and the morphological variation data of the experimental collection are also derived from photographs instead of remeasurements of the original replicas. Given the irreversible nature of archaeological excavations, digitized data, be it text, pictures, or videos, often become the sole evidence that is available for certain research questions. Yet, it has been widely acknowledged that the reuse of archaeological data has not received enough attention among researchers in our discipline (Faniel et al., 2018; Huggett, 2018; Moody et al., 2021). Among many reasons preventing archaeologists from reusing published and digitized data (Sobotkova, 2018), the lack of a standardized practice of and motivation for data sharing is a prominent one (Marwick & Birch, 2018). As stated in the method section, we addressed this issue by sharing the raw data and the code for generating the derived data on an open-access repository. Another major and legitimate concern of archaeological data reuse is their quality. In terms of this aspect, we do acknowledge the limitations of relying on photos when it comes to the more detailed technological analysis of stone artifacts, however, our paper shows that finding the appropriate research questions given the data available is key to revealing new novel insights into the studied topic. Moreover, we believe that this type of research has a strong contemporary relevance due to the continued influence of the COVID-19 on fieldwork-related travel and direct access to archaeological artifacts (Balandier et al., 2022; Ogundiran, 2021).

5 Conclusions

Regarding the two research questions we proposed in the beginning, our case study suggested that 1) we can delineate the effects of skill level and mental template through the multivariate analysis of morphometric data, where the former is associated with cross-sectional thinning while the latter is reflected in elongation and pointedness; 2) Training has an immediate effect of sharing a common mental template, but 90 hours of training is still not enough for novice to reach the level of expertise as reflected in modern experienced knappers, let alone the Boxgrove foragers. However, it should be noted that it is not our intention to construct a false binary framework and put these two factors as disconnected and opposite concepts. In the future, more robust experimental studies are needed to deepen our understanding of the relationship between skill acquisition and the morphological variability of bifaces as well as their implications for the biological and cultural evolution of the hominin lineages.

6 CRediT authorship contribution statement

Cheng Liu: Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Nada Khreisheh**: Investigation, Writing – review & editing. **Radu Iovita**: Resources, Writing – review & editing. **Dietrich Stout**: Conceptualization, Investigation, Resources, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing. **Justin Pargeter**: Conceptualization, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing.

7 Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

8 Acknowledgements

This work was supported by funding from the National Science Foundation of the USA (grants SMA-1328567 & DRL-1631563), the John Templeton Foundation (grant 47994), and the Emory University Research Council. The handaxe knapping skill acquisition experiment involved in this study was approved by Emory University's Internal Review Board (IRB study no: 00067237).

References

Balandier, C., Cipin, I., Hartenberger, B., & Islam, M. (2022). Archaeology in a pandemic: Four stories. *Near Eastern Archaeology*, *85*(1), 66–73. https://doi.org/10.1086/718201

Bamforth, D. B., & Finlay, N. (2008). Introduction: Archaeological approaches to lithic production skill and craft learning. *Journal of Archaeological Method and Theory*, *15*(1), 1–27. https://www.jstor.org/stable/40345992

Bayani, K. Y. T., Natraj, N., Khresdish, N., Pargeter, J., Stout, D., & Wheaton, L. A. (2021). Emergence of perceptuomotor relationships during paleolithic stone toolmaking learning: intersections of observation and practice. *Communications Biology*, *4*(1), 1–12. https://doi.org/10.1038/s42003-021-02768-w

Bello, S. M., Parfitt, S. A., & Stringer, C. (2009). Quantitative micromorphological analyses of cut marks produced by ancient and modern handaxes. *Journal of Archaeological Science*, *36*(9), 1869–1880. https://doi.org/10.1016/j.jas.2009.04.014

Bradley, B. A., & Sampson, C. G. (1986). *Analysis by replication of two acheuleian artefact assemblages from caddington, england* (G. Bailey & P. Callow, Eds.; pp. 29–46). Cambridge University Press.

Callahan, E. (1979). The basics of biface knapping in the eastern fluted point tradition: A manual for flintknappers and lithic analysts. *Archaeology of Eastern North America*, *7*(1), 1–180.

https://www.jstor.org/stable/40914177

Caruana, M. V. (2022). Extrapolating later acheulian handaxe reduction sequences in south africa: A case study from the cave of hearths and amanzi springs. *Lithic Technology*, *47*(1), 1–12. https://doi.org/10.1080/01977261.2021.1924452

Caruana, M. V. (2020). South African handaxes reloaded. *Journal of Archaeological Science: Reports*, *34*, 102649. https://doi.org/10.1016/j.jasrep.2020.102649

Caruana, M. V., & Herries, A. I. R. (2021). Modelling production mishaps in later Acheulian handaxes from the Area 1 excavation at Amanzi Springs (Eastern Cape, South Africa) and their effects on reduction and morphology. *Journal of Archaeological Science: Reports*, *39*, 103121. https://doi.org/10.1016/j.jasrep.2021.103121

Clark, J. D. (2001). *Variability in primary and secondary technologies of the later acheulian in africa* (S. Milliken & J. Cook, Eds.; p. 118). Oxbow Books.

Connolly, K., & Dalgleish, M. (1989). The emergence of a tool-using skill in infancy. *Developmental Psychology*, *25*(6), 894–912. https://doi.org/10.1037/0012-1649.25.6.894

Corbey, R. (2020). Baldwin effects in early stone tools. *Evolutionary Anthropology: Issues, News, and Reviews*, *29*(5), 237–244. https://doi.org/10.1002/evan.21864

Corbey, R., Jagich, A., Vaesen, K., & Collard, M. (2016). The acheulean handaxe: More like a bird's song than a beatles' tune? *Evolutionary Anthropology: Issues, News, and Reviews*, *25*(1), 6–19. https://doi.org/10.1002/evan.21467

Deetz, J. (1967). *Invitation to archaeology*. Natural History Press.

Eren, M. I., Roos, C. I., Story, B. A., von Cramon-Taubadel, N., & Lycett, S. J. (2014). The role of raw material differences in stone tool shape variation: an experimental assessment. *Journal of Archaeological Science*, *49*, 472–487. https://doi.org/10.1016/j.jas.2014.05.034

Faisal, A., Stout, D., Apel, J., & Bradley, B. (2010). The Manipulative Complexity of Lower Paleolithic Stone Toolmaking. *PLOS ONE*, *5*(11), e13718. https://doi.org/10.1371/journal.pone.0013718

Faniel, I. M., Austin, A., Kansa, E., Kansa, S. W., France, P., Jacobs, J., Boytner, R., & Yakel, E. (2018). Beyond the Archive: Bridging Data Creation and Reuse in Archaeology. *Advances in Archaeological Practice*, *6*(2), 105–116. https://doi.org/10.1017/aap.2018.2

García-Medrano, P., Ashton, N., Moncel, M.-H., & Ollé, A. (2020). The WEAP method: A new age in the analysis of the Acheulean handaxess. *Journal of Paleolithic Archaeology*, *3*(4). https://doi.org/10.1007/s41982-020-00054-5

García-Medrano, P., Maldonado-Garrido, E., Ashton, N., & Ollé, A. (2020). Objectifying processes: The use of geometric morphometrics and multivariate analyses on Acheulean tools. *Journal of Lithic Studies*, *7*(1). https://doi.org/10.2218/jls.4327

García-Medrano, P., Ollé, A., Ashton, N., & Roberts, M. B. (2019). The Mental Template in Handaxe Manufacture: New Insights into Acheulean Lithic Technological Behavior at Boxgrove, Sussex, UK. *Journal of Archaeological Method and Theory*, *26*(1), 396–422. https://doi.org/10.1007/s10816-018-9376-0

Herzlinger, G., Goren-Inbar, N., & Grosman, L. (2017). A new method for 3D geometric morphometric shape analysis: The case study of handaxe knapping skill. *Journal of Archaeological Science: Reports*, *14*, 163–173. https://doi.org/10.1016/j.jasrep.2017.05.013

Hillson, S. W., Parfitt, S. A., Bello, S. M., Roberts, M. B., & Stringer, C. B. (2010). Two hominin incisor teeth from the middle Pleistocene site of Boxgrove, Sussex, England. *Journal of Human Evolution*, *59*(5), 493–503. https://doi.org/10.1016/j.jhevol.2010.06.004

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022). People construct simplified mental representations to plan. *Nature*, *606*(7912), 129–136. https://doi.org/10.1038/s41586-022-04743-9

Hodgson, D. (2015). The symmetry of Acheulean handaxes and cognitive evolution. *Journal of Archaeological Science: Reports*, *2*, 204–208. https://doi.org/10.1016/j.jasrep.2015.02.002

Holmes, J. A., Atkinson, T., Fiona Darbyshire, D. P., Horne, D. J., Joordens, J., Roberts, M. B., Sinka, K. J., & Whittaker, J. E. (2010). Middle Pleistocene climate and hydrological environment at the Boxgrove hominin site (West Sussex, UK) from ostracod records. *Quaternary Science Reviews*, *29*(13), 1515–1527. https://doi.org/10.1016/j.quascirev.2009.02.024

Huggett, J. (2018). Reuse Remix Recycle: Repurposing Archaeological Digital Data. *Advances in Archaeological Practice*, *6*(2), 93–104. https://doi.org/10.1017/aap.2018.1

Hutchence, L., & Debackere, S. (2019). An evaluation of behaviours considered indicative of skill in handaxe manufacture. *LithicsThe Journal of the Lithic Studies Society*, *39*, 36.

Hutchence, L., & Scott, C. (2021). Is Acheulean Handaxe Shape the Result of Imposed 'Mental Templates' or Emergent in Manufacture? Dissolving the Dichotomy through Exploring 'Communities of Practice' at Boxgrove, UK. *Cambridge Archaeological Journal*, *31*(4), 675–686. https://doi.org/10.1017/S0959774321000251

Iovita, R., & McPherron, S. P. (2011). The handaxe reloaded: A morphometric reassessment of Acheulian and Middle Paleolithic handaxes. *Journal of Human Evolution*, *61*(1), 61–74. https://doi.org/10.1016/j.jhevol.2011.02.007

Iovita, R., Tuvi-Arad, I., Moncel, M.-H., Despriée, J., Voinchet, P., & Bahain, J.-J. (2017). High handaxe symmetry at the beginning of the European Acheulian: The data from la Noira (France) in context. *PLOS ONE*, *12*(5), e0177063. https://doi.org/10.1371/journal.pone.0177063

Isaac, G. L. (1986). *Foundation stones: Early artefacts as indicators of activities and abilities* (G. Bailey & P. Callow, Eds.; pp. 221–241). Cambridge University Press.

Kempe, M., Lycett, S., & Mesoudi, A. (2012). An experimental test of the accumulated copying error model of cultural mutation for Acheulean handaxe size. *PLOS ONE*, *7*(11), e48333. https://doi.org/10.1371/journal.pone.0048333

Key, A. J. M., & Lycett, S. J. (2017). Influence of Handaxe Size and Shape on Cutting Efficiency: A Large-Scale Experiment and Morphometric Analysis. *Journal of Archaeological Method and Theory*, *24*(2), 514–541. https://doi.org/10.1007/s10816-016-9276-0

Key, A. J. M., & Lycett, S. J. (2019). Biometric variables predict stone tool functional performance more effectively than tool-form attributes: a case study in handaxe loading capabilities. *Archaeometry*, *61*(3), 539–555. https://doi.org/10.1111/arcm.12439

Key, A. J. M., Proffitt, T., Stefani, E., & Lycett, S. J. (2016). Looking at handaxes from another angle: Assessing the ergonomic and functional importance of edge form in Acheulean bifaces. *Journal of Anthropological Archaeology*, *44*, 43–55. https://doi.org/10.1016/j.jaa.2016.08.002

Khreisheh, N. N., Davies, D., & Bradley, B. A. (2013). Extending Experimental Control: The Use of Porcelain in Flaked Stone Experimentation. *Advances in Archaeological Practice*, *1*(1), 38–46. https://doi.org/10.7183/2326-3768.1.1.37

Kohn, M., & Mithen, S. (1999). Handaxes: products of sexual selection? *Antiquity*, *73*(281), 518–526. https://doi.org/10.1017/S0003598X00065078

Kolhatkar, M. (2022). Skill in Stone Knapping: an Ecological Approach. *Journal of Archaeological Method and Theory*, *29*(1), 251–304. https://doi.org/10.1007/s10816-021-09521-x

Lycett, S. J., & Gowlett, J. A. J. (2008). On questions surrounding the acheulean 'tradition'. *World Archaeology*, *40*(3), 295–315. https://www.jstor.org/stable/40388215

Lycett, S. J., Schillinger, K., Eren, M. I., von Cramon-Taubadel, N., & Mesoudi, A. (2016). Factors affecting Acheulean handaxe variation: Experimental insights, microevolutionary processes, and macroevolutionary outcomes. *Quaternary International*, *411*, 386–401. https://doi.org/10.1016/j.quaint.2015.08.021

Lycett, S. J., von Cramon-Taubadel, N., & Foley, R. A. (2006). A crossbeam co-ordinate caliper for the morphometric analysis of lithic nuclei: a description, test and empirical examples of application. *Journal of Archaeological Science*, *33*(6), 847–861. https://doi.org/10.1016/j.jas.2005.10.014

Lyman, R. L., & O'Brien, M. J. (2004). A History of Normative Theory in Americanist Archaeology. *Journal of Archaeological Method and Theory*, *11*(4), 369–396. https://doi.org/10.1007/s10816-004-1420-6

Machin, A. J., Hosfield, R. T., & Mithen, S. J. (2007). Why are some handaxes symmetrical? Testing the influence of handaxe morphology on butchery effectiveness. *Journal of Archaeological Science*, *34*(6), 883–893. https://doi.org/10.1016/j.jas.2006.09.008

Marwick, B. (2017). Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation. *Journal of Archaeological Method and Theory*, *24*(2), 424–450. https://doi.org/10.1007/s10816-015-9272-9

Marwick, B., & Birch, S. E. P. (2018). A Standard for the Scholarly Citation of Archaeological Data as an Incentive to Data Sharing. *Advances in Archaeological Practice*, *6*(2), 125–143. https://doi.org/10.1017/aap.2018.3

Milks, A. (2019). Skills shortage: a critical evaluation of the use of human participants in early spear experiments. *EXARC Journal*, *2019*(2), 1–11. https://pdf.printfriendly.com/pdfs/make

Moody, B., Dye, T., May, K., Wright, H., & Buck, C. (2021). Digital chronological data reuse in archaeology: Three case studies with varying purposes and perspectives. *Journal of Archaeological Science: Reports*, *40*, 103188. https://doi.org/10.1016/j.jasrep.2021.103188

Nowell, A. (2002). Coincidental factors of handaxe morphology. *Behavioral and Brain Sciences*, *25*(3), 413–414. https://doi.org/10.1017/S0140525X02330073

Nowell, A., & White, M. (2010). *Growing up in the middle pleistocene: Life history strategies and their relationship to acheulian industries.* (A. Nowell & I. Davidson, Eds.; pp. 67–82). University Press of Colorado. http://www.upcolorado.com/book/Stone_Tools_and_the_Evolution_of_Human_Cognition_Paper

Ogundiran, A. (2021). Doing Archaeology in a Turbulent Time. *African Archaeological Review*, *38*(3), 397–401. https://doi.org/10.1007/s10437-021-09460-8

Pargeter, J., Khreisheh, N., Shea, J. J., & Stout, D. (2020). Knowledge vs. know-how? Dissecting the foundations of stone knapping skill. *Journal of Human Evolution*, *145*, 102807. https://doi.org/10.1016/j.jhevol.2020.102807

Pargeter, J., Khreisheh, N., & Stout, D. (2019). Understanding stone tool-making skill acquisition: Experimental methods and evolutionary implications. *Journal of Human Evolution*, *133*, 146–166. https://doi.org/10.1016/j.jhevol.2019.05.010

Pelcin, A. (1997). The Effect of Indentor Type on Flake Attributes: Evidence from a Controlled Experiment. *Journal of Archaeological Science*, *24*(7), 613–621. https://doi.org/10.1006/jasc.1996.0145

Pelegrin, J. (1993). *A framework for analysing prehistoric stone tool manufacture and a tentative application to some early stone industries* (pp. 302–317). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198522638.003.0018

Petraglia, M. D., & Korisettar, R. (Eds.). (1998). *Early human behaviour in global context: The rise and diversity of the lower palaeolithic record.* Routledge. https://doi.org/10.4324/9780203203279

Pope, M., Parfitt, S., & Roberts, M. (2020). *The horse butchery site 2020: A high-resolution record of lower palaeolithic hominin behviour at boxgrove, UK.* SpoilHeap Publications.

Richerson, P. J., & Boyd, R. (2005). *Not By Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press.

Roberts, M. B., & Parfitt, S. A. (1998). *Boxgrove: A middle pleistocene hominid site at eartham quarry, boxgrove, west sussex*. English Heritage.

Roberts, M. B., & Pope, M. (2009). *The archaeological and sedimentary records from boxgrove and slindon* (R. M. Briant, M. R. Bates, R. Hosfield, & F. Wenban-Smith, Eds.; pp. 96–122). Quaternary Research Association.

Roe, D. A. (1969). British Lower and Middle Palaeolithic Handaxe Groups*. *Proceedings of the Prehistoric Society*, *34*, 1–82. https://doi.org/10.1017/S0079497X00013840

Roux, V., Bril, B., & Dietrich, G. (1995). Skills and learning difficulties involved in stone knapping: The case of stone-bead knapping in khambhat, india. *World Archaeology*, *27*(1), 63–87. https://doi.org/10.1080/00438243.1995.9980293

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, *18*(1), 529. https://doi.org/10.1186/s12859-017-1934-z

Schick, K. D., & Toth, N. P. (1993). *Making Silent Stones Speak: Human Evolution And The Dawn Of Technology*. Simon; Schuster.

Sharon, G. (2008). The impact of raw material on Acheulian large flake production. *Journal of Archaeological Science*, *35*(5), 1329–1344. https://doi.org/10.1016/j.jas.2007.09.004

Sharon, G., Alperson-Afil, N., & Goren-Inbar, N. (2011). Cultural conservatism and variability in the Acheulian sequence of Gesher Benot Ya'aqov. *Journal of Human Evolution*, *60*(4), 387–397. https://doi.org/10.1016/j.jhevol.2009.11.012

Shipton, C., & Clarkson, C. (2015). Handaxe reduction and its influence on shape: An experimental test and archaeological case study. *Journal of Archaeological Science: Reports*, *3*, 408–419. https://doi.org/10.1016/j.jasrep.2015.06.029

Shipton, C., Clarkson, C., Pal, J. N., Jones, S. C., Roberts, R. G., Harris, C., Gupta, M. C., Ditchfield, P. W., & Petraglia, M. D. (2013). Generativity, hierarchical action and recursion in the technology of the Acheulean to Middle Palaeolithic transition: A perspective from Patpara, the Son Valley, India. *Journal of Human Evolution*, *65*(2), 93–108. https://doi.org/10.1016/j.jhevol.2013.03.007

Shipton, C., Petraglia, M. D., & Paddayya, K. (2009). Stone tool experiments and reduction methods at the Acheulean site of Isampur Quarry, India. *Antiquity*, *83*(321), 769–785. https://doi.org/10.1017/S0003598X00098987

Smith, G. M. (2013). Taphonomic resolution and hominin subsistence behaviour in the Lower Palaeolithic: differing data scales and interpretive frameworks at Boxgrove and Swanscombe (UK). *Journal of Archaeological Science*, *40*(10), 3754–3767. https://doi.org/10.1016/j.jas.2013.05.002

Smith, G. M. (2012). Hominin-carnivore interaction at the Lower Palaeolithic site of Boxgrove, UK. *Journal of taphonomy*, *10*(3-4), 373–394. https://dialnet.unirioja.es/servlet/articulo?codigo=5002455

Sobotkova, A. (2018). Sociotechnical Obstacles to Archaeological Data Reuse. *Advances in Archaeological Practice*, *6*(2), 117–124. https://doi.org/10.1017/aap.2017.37

Sterelny, K. (2004). A review of Evolution and learning: the Baldwin effect reconsidered edited by Bruce Weber and David Depew. *Evolution & Development*, *6*(4), 295–300. https://doi.org/10.1111/j.1525-142X.2004.04035.x

Stout, D. (2002). Skill and cognition in stone tool production: An ethnographic case study from irian jaya. *Current Anthropology*, *43*(5), 693–722. https://doi.org/10.1086/342638

Stout, D., Apel, J., Commander, J., & Roberts, M. (2014). Late Acheulean technology and cognition at Boxgrove, UK. *Journal of Archaeological Science*, *41*, 576–590. https://doi.org/10.1016/j.jas.2013.10.001

Stout, D., Passingham, R., Frith, C., Apel, J., & Chaminade, T. (2011). Technology, expertise and social cognition in human evolution. *European Journal of Neuroscience*, *33*(7), 1328–1338. https://doi.org/10.1111/j.1460-9568.2011.07619.x

Wenban-Smith, F. (2004). Handaxe typology and Lower Palaeolithic cultural development: ficrons, cleavers and two giant handaxes from Cuxton. *Lithics*, *25*, 11–21. https://eprints.soton.ac.uk/41481/

Wenban-Smith, F., Gamble, C., & Apsimon, A. (2000). The Lower Palaeolithic Site at Red Barns, Portchester, Hampshire: Bifacial Technology, Raw Material Quality, and the Organisation of Archaic Behaviour. *Proceedings of the Prehistoric Society*, *66*, 209–255. https://doi.org/10.1017/S0079497X0000181X

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.

White, M. (1998). On the Significance of Acheulean Biface Variability in Southern Britain. *Proceedings of the Prehistoric Society*, *64*, 15–44. https://doi.org/10.1017/S0079497X00002164

White, M. (1995). Raw materials and biface variability in southern britain: A preliminary examination. *LithicsThe Journal of the Lithic Studies Society*, *15*, 1–20.

White, M., & Foulds, F. (2018). Symmetry is its own reward: on the character and significance of Acheulean handaxe symmetry in the Middle Pleistocene. *Antiquity*, *92*(362), 304–319. https://doi.org/10.15184/aqy.2018.35