Received: 15 December 2018 | Accepted: 25 January 2019
DOI: 10.1002/evan.21772

NEWS

Evolutionary Anthropology ISSUES, NEWS, AND REVIEWS WILEY

Comparative analysis of Middle Stone Age artifacts in Africa (CoMSAfrica)

1 | INTRODUCTION

Spatial and temporal variation among African Middle Stone Age (MSA) archeological assemblages provide essential cultural and behavioral data for understanding the origin, evolution, diversification, and dispersal of *Homo sapiens*—and, possibly, interactions with other hominin taxa.[1,2] However, incorporating archeological data into a robust framework suited to replicable, quantitative analyses that can be integrated with observations drawn from studies of the human genome, hominin morphology, and paleoenvironmental contexts requires the development of a unified comparative approach and shared units of analysis. Lithic (stone) artifacts provide the fundamental source of information for continental-scale comparisons of past hominin behavior because they quantitatively dominate the Paleolithic record, and unlike organic artifacts made of bone or shell, they are preserved in a larger variety of depositional settings. However, attempts to integrate African MSA lithic data from different periods or regions have suffered from divergent research traditions among archeologists that employ incompatible approaches, definitions, and data collection methods. Communication among analysts is further constrained by the presence of varied theoretical and methodological schools, including analytical grammars that may represent distinct ways of viewing, describing, measuring, and interpreting the world (i.e., attribute analysis vs. *chaîne opératoire*). These issues are further exacerbated by differences in geography, geology, ecology, and research intensity between different parts of Africa. Archeologists across Africa thus lack a common, intersubjective and transparent system for lithic analysis, with currently few shared basic definitions or protocols of measurements. Yet, objectivity and replicability are two functional requirements of science.

2 | WORKSHOP OVERVIEW

The workshop “Comparative Analysis of Middle Stone Age Artefacts in Africa (CoMSAfrica)” aimed to reflect upon a common and replicable analytical framework, as well as proposing concrete solutions for its implementation. It builds on previous efforts to standardize pan-African comparisons which focused on higher taxonomic entities, specific categories of stone artifacts, or individual regions.[3,4] Organized by C. Tryon and M. Will, the workshop brought together 12 international scholars (see author list) working in different periods and regions of Africa, with varied methodological backgrounds. The workshop was held between November 5th and 6th 2018 at Harvard University (USA), and funded through the Accelerator Workshop Program of the Radcliffe Institute for Advanced Study (https://projects.iq.harvard.edu/comsafrica). The meeting included short introductory presentations by all participants followed by a series of more focused roundtable discussions to define the main problems and issues confronting comparative lithic analyses of African MSA assemblages. The final goal was the development of a unified analytical approach. As a two-day workshop was obviously insufficient to solve problems of such magnitude, the final discussion focused on outlining a working model and roadmap for future meetings and collaborations through the CoMSAfrica network.

3 | OUTCOMES

Comparability and replicability require explicit and unambiguous definitions shared across researchers. Initial group discussions clarified the need to focus on simple, individual artifact attributes, such as length, where interobserver error and appeals to expert knowledge (i.e., difference in the experience of analysts) could be minimized, or at least better constrained, reflecting an underlying trade-off between analytic complexity and replicability. It also became clear that definitions of even fundamental artifact classes such as cores and retouched tools were strongly affected by a researcher's experience and methodological background. Here, the typical application of compound definitions—those in which a given class is identified as the product of multiple independent attributes that often have nonbinary states—increased the probability of interanalyst divergence. While considered tractable, it was accepted that the two days available were insufficient to resolve these complexities, or the dependent issues of core and tool typologies, and these were set aside for subsequent workshops. The broader implications of these problems should not, however, be overlooked: continent-wide, or even regional meta-analyses of existing lithic data sets are confounded to a considerable degree by interanalyst variance (e.g., Tryon and Faith[5]).

Following these realizations, the workshop focused on the fundamental issue of interanalyst variance as a starting point, in particular by giving attention to the recording of unretouched flakes only. Attributes used to describe unretouched flakes appeared to show higher levels of interanalyst agreement and lesser involvement of expert knowledge, and thus potentially the highest comparability among the workshop participants. Unretouched flakes are also the most frequent category of finds in MSA lithic assemblages. They carry important information about production methods and techniques, as well as information on reduction sequence and intensity, and have featured prominently in recent quantitative studies of temporal and spatial technological variability among early *H. sapiens* populations (e.g., Tostevin[6]; Scerri et al.[7]).

Workshop participants developed a working minimum set of about 40 attributes considered useful for reliable comparative analyses, using previous efforts to provide a standardized list of attributes suitable for inter-site studies (e.g., Wilkins et al.[8]) as a starting point. We discovered that databases specifically designed for a single African region do not work equally well for all parts of the continent. Discussions about attribute definition, coding, and collection method are ongoing, a process that revealed substantial variation in the way that even seemingly basic attributes such as flake thickness, proportion of cortex, and dorsal scar patterns were defined, measured, and assessed. One of the clearest outcomes was the need to first develop a robust program to evaluate interobserver variability in each of the chosen variables to establish baseline uncertainty estimates before moving to collect, compare, and contrast data aggregated from MSA sites across Africa. Otherwise, we might be comparing differences in the behavior of contemporary archeologists rather than that of ancient hominins. We have initiated a round-robin replication study in which all workshop participants analyze an experimentally generated flake assemblage, using our 40-attribute list to evaluate and quantify interobserver error on each measure, similar to interlaboratory studies done for radiocarbon dating and tephra correlation.[9,10]

4 | FUTURE DIRECTIONS

4.1 | How to continue the CoMSAfrica program?

All participants agreed on the need to create an enduring research network and communication platform for tackling the problem of comparative MSA lithic studies, and to continue the work started by the inaugural CoMSAfrica workshop. As seen in the current "replication crisis" affecting many of the sciences,[11] the problem of replicability and comparability between researchers is not unique to archeology, nor is it something that will go away soon or be solved by individuals alone. The workshop participants decided to strengthen within-group research cooperation, to establish a digital communication platform (i.e., Open Science Framework; Slack), to hold future workshops to develop the initial approach summarized here, and to present the results at professional meetings held in Africa and elsewhere. The long-term aim is to open up the network to all researchers working on the MSA of Africa, once basic issues with the recording framework and database are resolved.

4.2 | How to develop the method and database?

Although one outcome of the workshop was a working list of attributes for unretouched flakes that we believe will be useful for large-scale comparisons among MSA assemblages, the reliability and replicability of many of these attributes remains to be demonstrated. Individual studies have examined interobserver variance for various kinds and aspects of lithic artifacts, such as the identification of Levallois flakes, raw material types, and measurements.[12–14] Differentiating between individual attributes that can and that cannot be reliably compared between analysts has to precede any future application of the recording framework itself. Balancing replicability with expert knowledge—which is required for any in-depth analysis of stone tools—constitutes another key concern. The planned replication study will focus on these issues and specify levels of interobserver variance that are considered to be acceptable from the outset as part of the study design. Another issue to be assessed is the concern that the attributes chosen be applicable to all lithic raw materials, as the material properties of quartz in particular have often caused analytical problems, yet MSA assemblages made from this material dominate many regions (e.g., Equatorial Africa). In terms of the envisioned database itself, we are currently exploring multiple digital and open access platforms.

4.3 | How do we plan to apply these data?

Ultimately, the goal of CoMSAfrica is to generate data that can be used for comparative studies of African MSA lithic artifacts to answer behavioral questions about the past on multiple temporal and geographical scales. This will require a combination of existing data sets by the participants using attributes that are found to be most comparable between researchers in our replication tests. Only with an approach like this can we reliably perform continental-scale comparisons and assess the temporal and spatial variability among African MSA sites. Such replicable archeological data will provide the basis for testing models of cultural change, spatial isolation, and dispersal with the goal of ultimately integrating this information with evidence from hominin morphology, genetics, and paleoenvironmental reconstructions.

ACKNOWLEDGMENTS

The authors thank Sean O'Donnell, Kristen Osborne, Wendy Frohlich, and Maura Madden from Harvard University's Radcliffe Institute for Advanced Study for making this workshop both fun and possible. This is CoMSAfrica publication number 1.

ORCID

Manuel Will https://orcid.org/0000-0001-8116-2543
Christian Tryon https://orcid.org/0000-0002-2354-3273
Justin Pargeter https://orcid.org/0000-0001-6536-8568
Alice Leplongeon https://orcid.org/0000-0001-8464-8063
Huw S. Groucutt https://orcid.org/0000-0002-9111-1720
Alison S. Brooks https://orcid.org/0000-0002-7978-8747

Manuel Will[1], Christian Tryon[2], Matthew Shaw[3]
Eleanor M. L. Scerri[4,5], Kathryn Ranhorn[2], Justin Pargeter[6,7],
Jessica McNeil[2], Alex Mackay[3], Alice Leplongeon[8,9],
Huw S. Groucutt[4,5,10], Katja Douze[11], Alison S. Brooks[12,13]

[1]Department of Early Prehistory and Quaternary Ecology, University of Tübingen, Tübingen, Germany
[2]Department of Anthropology, Peabody Museum of Archaeology and Ethnology, Harvard University, Cambridge, Massachusetts
[3]Centre for Archaeological Science, University of Wollongong, Wollongong, New South Wales, Australia
[4]Max Planck Institute for the Science of Human History, Jena, Germany
[5]School of Archaeology, University of Oxford, Oxford, United Kingdom
[6]Department of Anthropology, Emory University, Atlanta, Georgia
[7]Centre for Anthropological Research & Department of Anthropology and Development Studies, University of Johannesburg, Johannesburg, South Africa
[8]Institute of Advanced Studies and Department of Cultural Heritage, University of Bologna, Bologna, Italy
[9]UMR CNRS 7194, Département Homme et Environnement, Muséum national d'Histoire naturelle, Université de Perpignan Via Domitia, Alliance Sorbonne Université, Paris, France
[10]Max Planck Institute for Chemical Ecology, Jena, Germany
[11]Department of Genetics and Evolution, Anthropology Unit, Archaeology and Population in Africa, University of Geneva, Geneva, Switzerland
[12]Center for the Advanced Study of Human Paleobiology, Department of Anthropology, George Washington University, Washington, District of Columbia
[13]Human Origins Program, National Museum of Natural History, Smithsonian Institution, Washington, District of Columbia

Correspondence
Manuel Will
Email: manuel.will@uni-tuebingen.de

REFERENCES

[1] Groucutt HS, Petraglia MD, Bailey G, et al. 2015. Rethinking the dispersal of *Homo sapiens* out of Africa. Evol Anthropol 24: 149–164.

[2] Scerri EM, Thomas MG, Manica A, Gunz P, Stock JT, et al. 2018. Did our species evolve in subdivided populations across Africa, and why does it matter? Trends Ecol Evol 33:582–594.

[3] Clark JD, Cole GH, Isaac GL, Kleindienst MR. 1966. Precision and definition in African archaeology. S Afr Archaeol Bull 21: 114–121.

[4] Conard NJ, Soressi M, Parkington JE, Wurz S, Yates R. 2004. A unified lithic taxonomy based on patterns of core reduction. S Afr Archaeol Bull 59:12–16.

[5] Tryon CA, Faith JT. 2013. Variability in the Middle Stone Age of eastern Africa. Curr Anthropol 54:S234–S254.

[6] Tostevin GB. 2012. Seeing lithics: a middle-range theory for testing cultural transmission in the Pleistocene, Oxford: Oxbow Books.

[7] Scerri EM, Drake NA, Jennings R, Groucutt HS. 2014. Earliest evidence for the structure of *Homo sapiens* populations in Africa. Quat Sci Rev 101:207–216.

[8] Wilkins J, Brown KS, Oestmo S, et al. 2017. Lithic technological responses to late Pleistocene glacial cycling at pinnacle point site 5-6, South Africa. PLoS One 12:e0174051.

[9] Scott EM, Naysmith P, Cook GT. 2018. Why do we need ^{14}C inter-comparisons? The Glasgow-^{14}C inter-comparison series, a reflection over 30 years. Quat Geochronol 43:72–82.

[10] Kuehn SC, Froese DG, Shane PA, INTAV Intercomparison Participants. 2011. The INTAV intercomparison of electron-beam microanalysis of glass by tephrochronology laboratories: results and recommendations. Quat Int 246(1–2):19–47.

[11] Baker M. 2016. Is there a reproducibility crisis? Nature 533(7604): 452–455.

[12] Perpère M. 1986. Apport de la typométrie à la définition des éclats Levallois: l'exemple d'Ault. Bull Soc Préhist Fr 83(4): 115–118.

[13] Calogero B. 1992. Lithic misidentification. Man in the Northeast 43: 87–90.

[14] Dibble HL, Bernard MC. 1980. A comparative study of basic edge angle measurement techniques. Am Antiq 45:857–865.