# Evaluating claims for an early peopling of the Americas: experimental design and the Cerutti Mastodon site

Matthew Magnani[1,*], Dalyn Grindle[1], Sarah Loomis[1], Alexander M. Kim[1,2], Vera Egbers[1,3], Jon Clindaniel[1], Alexis Hartford[1], Eric Johnson[1], Sadie Weber[1] & Wade Campbell[1]


*In a 2017 article, Holen and colleagues reported evidence for a 130 000-year-old archaeological site in California. Acceptance of the site would overturn current understanding of global human migrations. The authors here consider Holen* et al.*'s conclusions through critical evaluation of their replicative experiments. Drawing on best practice in experimental archaeology, and paying particular attention to the authors' chain of inference, Magnani* et al. *suggest that to argue convincingly for an early human presence at the Cerutti Mastodon site, Holen* et al. *must improve their analogical foundations, test alternative hypotheses, increase experimental control and quantify their results.*

*Keywords:* North America, Cerutti Mastodon, experimental archaeology, human migration

Debate

## Introduction

Current archaeological and genetic evidence suggests that Indigenous peoples in the Americas are descended from those of ancient Siberia, with founding populations separating from ancient North Asians *c.* 25 000–16 000 years ago (Raghavan *et al.* 2015; Skoglund & Reich

[1] *Harvard University, Department of Anthropology, 11 Divinity Avenue, Cambridge, MA 02138, USA*
[2] *Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, New Research Building, Room 260, Boston, MA 02115, USA*
[3] *Free University of Berlin, Department of Near Eastern Archaeology, Fabeckstraße 23–25, 14195 Berlin, Germany*
* *Author for correspondence (Email: matthewmagnani@g.harvard.edu)*


ANTIQUITY 93 369 (2019): 789–795 https://doi.org/10.15184/aqy.2019.14

2016; Moreno-Mayar *et al.* 2018; Potter *et al.* 2018). While the arrival and dispersal routes remain contentious (Pedersen *et al.* 2016; Lesnek *et al.* 2018), it is probable that, after passing south of the ice sheets, humans quickly dispersed across North and South America (Cinq-Mars 1978; Dillehay *et al.* 2008; Goebel *et al.* 2008; Graf *et al.* 2014). Against this background, recent work by Holen *et al.* (2017) reports new evidence for stone tool use at the 130 000-year-old Cerutti Mastodon site in California. The authors employ a series of experiments to support their argument that the lithic and faunal remains at the site are anthropogenically modified. Their findings, if accepted, would push the peopling of the Americas back by over 100 000 years, thereby rewriting the history of global human migrations.

Such an exceptional claim has invited heavy criticism. Disputing Holen *et al.* (2017), scholars have suggested alternative taphonomic explanations for the formation of the Cerutti Mastodon site. These include comparing the Cerutti assemblage to both anthropogenic and non-anthropogenic sites, questioning the absence of features generally associated with the presence of humans, and the possibility that recent disturbances may have altered the remains (Braje *et al.* 2017; Haynes 2017, 2018; Ferraro *et al.* 2018). Holen *et al.* (2018a–c) and their supporters (e.g. Boëda *et al.* 2017; Gruhn 2018) have responded to these criticisms, contending that alternative arguments fail to explain the features of the Cerutti Mastodon site.

While these previous responses to Holen *et al.* provide viable criticisms, the experimental design of the project used to support the archaeological claims has not been examined. Rather than reiterate the ongoing debate concerning the archaeological nature of the Cerutti site itself, here we examine the experimental data used to buttress Holen *et al.*'s argument. We raise questions about the researchers' analogical foundations, the lack of alternative hypotheses tested, as well as experimental control and the quantification of data. We conclude with recommendations for future experimentation intended to provide more robust evidence, through which Holen *et al.*'s (2017) claims may be more effectively evaluated by the archaeological community.

## Hypothesis testing and analogical reasoning in experimental archaeology

The first step in the development of an archaeological experiment entails hypothesis creation. The method by which one generates ideas should be clearly stated, as the interpretation of an experiment's results is directly dependent on the authors' underlying assumptions (Domínguez-Rodrigo 2008). The proposed hypothesis must be formulated in a falsifiable cause-and-effect scheme that consequently implies an opposing null hypothesis.

We follow the current best practice for archaeological experimentation proposed by scholars including Lin *et al.* (2017) and Eren *et al.* (2016). To the extent that Holen *et al.*'s logic of hypothesis construction is transparent or reproducible, we find it to be problematic. Instead of setting up the experiment with two opposing hypotheses, the authors have conducted their experiments and summarily confirmed that the observed patterns match those from the Cerutti site. While the authors do acknowledge alternatives, they do not evaluate the other viable explanations for the observed patterns (see also Braje *et al.* 2017, and the rebuttal by Holen *et al.* 2018c). In order to suggest that the bones were, in fact, modified by humans, Holen and





https://doi.org/10.15184/aqy.2019.14 Published online by Cambridge University Press

colleagues must also quantitatively compare lithic and faunal remains—modified via other taphonomic processes—against their experimental dataset.

Beyond hypothesis construction, it is critical to justify the analogues that scaffold the experiment. In their 2017 publication, the authors focus much of their attention on the use of a hafted hammer stone. The origin of this analogy should be explained—beyond a passing reference to late prehistoric Plains peoples (see Holen *et al.*'s 2017 supplementary information 5, p. 22). The replication of patterns from the La Sena Mammoth site (see also Holen *et al.*'s 2017 supplementary information 5, p. 22) represents an uncontextualised comparison that appears only tangentially related to the Cerutti Mastodon site—and nearly 100 000 years removed. The experiment should be set up so that the variables are as close as possible to the observation that one is attempting to replicate. In this case, the experiments presented are not only too variable and inconsistent for any meaningful assessment of hypotheses, but are also built on foundations that are insufficiently comparable to the archaeological context in question.

## Experimental control

When designing and implementing an experiment, it is important to exert control over variables at a level justified according to the experimental design (for a discussion on experimental validity, see Mesoudi 2011, and for an illustration of its spectrum, see Eren *et al.* 2016). This ensures that the observed results reflect manipulation of those variables associated with the hypotheses, rather than the interaction of nuisance variables (*sensu* Lin *et al.* 2017, but see also a discussion of the role of randomisation in mitigating nuisance variables on page 14 of the same manuscript). Based on a review of Holen *et al.*'s experimental design, we conclude that there was substantial variation not only in the raw materials used, but also the ways in which the experiments were executed; these variables detract substantially from the arguments made by the authors.

In reporting their study, Holen *et al.* (2017) list the raw materials and mass of individual tools from the Cerutti site, as well as their experimental hammer stones and anvils. While the individual archaeological hammer stones range in mass from 7.6–14.45kg (without refits; up to 18.25kg with refits), and are made of andesite and pegmatite, the experimental hammer stones range from 1.7–14.7kg and are of andesite or granite. Similar differences characterise the anvils. The variation in raw material, size and mass make it difficult to compare the results of the experiments with confidence. In turn, extrapolating to the archaeological record and the Cerutti assemblage becomes more tenuous.

Acknowledging the rarity of elephant remains with which to conduct such experiments, we recommend that scarce elephant bones not be used to confirm the hypothesis as presented by the authors, but rather function as benchmarks against which more refined experiments may be compared. For instance, after establishing whether or not the different sizes of hammer stones can be used effectively in the processing of large proboscidean bones, it would then be productive to establish a more robust experimental procedure using the bones of cattle or other large mammals.

Holen *et al.*'s four reported experiments were executed inconsistently. Of these experiments, two were conducted on elephant bones, while the other two were undertaken



Debate



https://doi.org/10.15184/aqy.2019.14 Published online by Cambridge University Press

using cattle and/or kangaroo bones. In the first case, the elephant bones were propped up on a wooden block and struck with a hafted stone. In a second experiment conducted on an elephant bone, a larger hammer stone was used after the failure of another hafted implement. The specifics of the subsequent lines of experimentation, carried out with cattle and kangaroo bones, are less clear, and require more detailed reporting to allow meaningful evaluation or repetition. We suggest future iterations of this experimentation begin by employing both hammers and anvils of approximately the same dimensions, composition and size. These should be randomly assigned to treatment groups, in order to create results that may be compared with one another and with outgroups of assemblages created through the testing of alternative hypotheses.

Once the raw materials used in the experimentation are held constant, it would then be beneficial to ensure that the ways in which the raw materials are made to interact (e.g. hammer stone on bone impacts) are also systematised. In reviewing the highly controlled experiments of Dibble and Rezek (2009), and work by Magnani *et al.* (2014), we understand that strict laboratory conditions are not always attainable, or desirable, in experimental archaeology. At a minimum, however, we suggest ensuring that the bones be struck and supported as consistently as possible. Ideally, the force would be regular and dealt from similar angles, a set number of times. The reported experiments make no apparent effort to control for any of these variables, casting doubt on the internal validity of the results (*sensu* Lin *et al.* 2017).

## Quantification of results

The experiments reported by Holen *et al.* (2017) are entirely qualitative, producing a good initial pilot study and the basis for designing further archaeological experiments, but limited inferential value (Lin *et al.* 2017: 680). When designing replication experiments, quantification is a basic requirement to enable the comparison of excavated artefacts with experimental results using standard statistical methods (Eren *et al.* 2016: 108). Quantification allows researchers to assess alternative hypotheses explicitly. Qualitative visual comparisons are insufficient to draw meaningful conclusions regarding differences in the morphological characteristics of bone and stone tool assemblages. Fortunately, there are numerous options for Holen and colleagues to consider. For stone tool morphometrics, for example, Caruana *et al.* (2014) suggest the use of digital elevation models in ArcGIS to quantify the shape of pitting on stone artefacts ($\frac{perimeter^2}{volume}$), in order to distinguish anthropogenic traces (whether ancient or modern) from natural characteristics of river cobbles. The authors may further wish to evaluate the morphologies of the bone flakes and impact notches from the Cerutti site, in relation to experimental datasets, employing both dimensional measurements standard in the assessment of anthropogenic bone modifications (Galán *et al.* 2009), as well as more recently developed geometric morphometric techniques (Yravedra *et al.* 2018). The use of any or all of these quantitative approaches would better enable the archaeological community to evaluate the authors' findings.

Finally, quantitative measures depend on assumptions of a sufficiently large experimental sample to ensure an acceptable level of statistical confidence (Drennan 2009: 126). In order to assess the proper sample size needed for an experiment, it is standard across experimental fields for researchers to perform a power analysis: a measure of how well a particular statistical test can distinguish a difference between groups, if indeed there is one (Cohen 1992: 100). After establishing appropriate controls on the raw materials and other experimental parameters, we therefore recommend that the authors employ such an analysis, expanding their experimental sample size appropriately to facilitate meaningful evaluation of their data.

## Conclusions

For over a century, anthropological enquiry into the peopling of the Americas has continually expanded to accommodate new archaeological evidence and genetic data. The Cerutti Mastodon site, as reported by Holen and colleagues, joins the ranks of archaeological discoveries that have challenged established narratives about global human migrations. Additional evidence and experiments, however, are required for the interpretation of this site to gain widespread acceptance by the archaeological community, and to identify meaningfully the presence of humans in the Americas over 100 000 years ago. While the genetics of present-day and ancient Native Americans do not indicate any obvious contribution from the authors' hypothesised early entrants into the Americas (Sankararaman *et al.* 2016; Browning *et al.* 2018; Moreno-Mayar *et al.* 2018), the lack of such a contribution does not refute their contention that Cerutti is an early human archaeological site. There are numerous cases of genetic discontinuity between present-day groups in many parts of the world and the earliest-documented inhabitants of those areas. For example, the oldest incontrovertible modern humans in Northern Eurasia—represented by the ancient genome of the 45 000-year-old Ust'-Ishim femur from West Siberia (Fu *et al.* 2014)—were not clearly ancestral to any present-day groups. Such may also be the case for the first humans in the Americas.

To be clear, we are not opposed categorically to the demonstration of an early human presence in the Americas. We do, however, expect that significant assertions in support of such a case, including those made by Holen and colleagues, must be built on a solid foundation, following current best practice in the field of experimental archaeology. Central to the argument of Holen *et al.* is their qualitative dataset derived from experimentation. In effect, the experiments used to support their 2017 article function primarily as a series of pilot studies. While such studies are an important part of the discipline of experimental archaeology, and may suggest preliminary relationships between variables, they are not sufficient in their own right to overturn decades of archaeological research on human migration.

Moving forward, we argue that analysis of the Cerutti site demands a more robust experimental archaeological study, which includes shoring up the analogical reasoning, controlling experimental parameters, increasing sample size and the quantification of results. These constitute the essential first steps that will allow for a more thorough evaluation of the exceptional claims made—as well as the dismissal of the alternative explanations offered—for the Cerutti site. We thank Holen and his co-authors for making extensive supplementary data available for their project, including videos and three-dimensional models, which will prove beneficial for further evaluation of the Cerutti site. Although our piece is a critical assessment, it is the availability of these data that has made it possible to evaluate the authors' interpretations, to draw conclusions and to make our recommendations.







https://doi.org/10.15184/aqy.2019.14 Published online by Cambridge University Press

## References


Boëda, E., C. Griggo & C. Lahaye. 2017. The Cerutti Mastodon site: archaeological or paleontological? *PaleoAmerica* 3: 193–95. https://doi.org/10.1080/20555563.2017.1338006

Braje, T., T. Dillehay, R. Kelly, R. Klein, D. Meltzer & T. Rick. 2017. Were hominins in California ~130,000 years ago? *PaleoAmerica* 3: 200–202. https://doi.org/10.1080/20555563.2017.1348091

Browning, S., B. Browning, Y. Zhou, S. Tucci & J. Akey. 2018. Analysis of human sequence data reveals two pulses of archaic Denisovan admixture. *Cell* 173: 53–61. https://doi.org/10.1016/j.cell.2018.02.031

Caruana, M., S. Varvalho, D. Braun, D. Presnyakova, M. Haslam, W. Archer, R. Bobe & J.W.K. Harris. 2014. Quantifying traces of tool use: a novel morphometric analysis of damage patterns on percussive tools. *PLoS ONE* 9: 1–18. https://doi.org/10.1371/journal.pone.0113856

Cinq-Mars, J. 1978. Bluefish Cave I: a Late Pleistocene Eastern Beringian cave deposit in the northern Yukon. *Canadian Journal of Anthropology* 3: 1–32.

Cohen, J. 1992. Statistical power analysis. *Current Directions in Psychological Science* 1: 98–101. https://doi.org/10.1111/1467-8721.ep10768783

Dibble, H.L. & Z. Rezek. 2009. Introducing a new experimental design for controlled studies of flake formation: results for exterior platform angle, platform depth, angle of blow, velocity, and force. *Journal of Archaeological Science* 36: 1945–54. https://doi.org/10.1016/j.jas.2009.05.004

Dillehay, T., C. Ramírez, M. Pino, M. Collins, J. Collins & J. Pino-Navarro. 2008. Monte Verde: seaweed, food, medicine, and the peopling of South America. *Science* 320: 784–86. https://doi.org/10.1126/science.1156533

Domínguez-Rodrigo, M. 2008. Conceptual premises in experimental design and their bearing on the use of analogy: an example from experiments on cut marks. *World Archaeology* 40: 67–82. https://doi.org/10.1080/00438240701843629

Drennan, R. 2009. *Statistics for archaeologists: a common sense approach*. New York: Springer. https://doi.org/10.1007/978-1-4419-0413-3

Eren, M., S. Lycett, R. Patten, B. Buchanan, J. Pargeter & M. O'Brien. 2016. Test, model, and method validation: the role of experimental stone artifact replication in hypothesis-driven archaeology. *Ethnoarchaeology* 8: 103–36. https://doi.org/10.1080/19442890.2016.1213972

Ferraro, J., K. Binetti, L. Weiest, D. Esker, L. Baker & S. Forman. 2018. Contesting early archaeology in California. *Nature* 554: E1–2. https://doi.org/10.1038/nature25165

Fu, Q. *et al.* 2014. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* 514: 445–49. https://doi.org/10.1038/nature13810

Galán, A., M. Rodríguez, S. De Juana & M. Domínguez-Rodrigo. 2009. A new experimental study on percussion marks and notches and their bearing on the interpretation of hammerstone-broken faunal assemblages. *Journal of Archaeological Science* 36: 776–84. https://doi.org/10.1016/j.ja

Goebel, T., M. Waters & D. O'Rourke. 2008. The Late Pleistocene dispersal of modern humans in the Americas. *Science* 319: 1497–502. https://doi.org/10.1126/science.1153569

Graf, K., C. Ketron & M. Water (ed.). 2014. *Paleoamerican odyssey*. College Station: Texas A&M University Consortium Press.

Gruhn, R. 2018. Observations concerning the Cerutti Mastodon site. *PaleoAmerica* 4: 101–102. https://doi.org/10.1080/20555563.2018.1467192

Haynes, G. 2017. The Cerutti mastodon. *PaleoAmerica* 3: 196–99. https://doi.org/10.1080/20555563.2017.1330103

– 2018. Reply to Holen *et al.* regarding the Cerutti mastodon. *PaleoAmerica* 4: 99–100. https://doi.org/10.1080/20555563.2018.1460562

Holen, S.R., T.A. Deméré, D.C. Fisher, R. Fullagar, J.B. Paces, G.T. Jefferson, J.M. Beeton, R.A. Cerutti, A.N. Rountrey, L. Vescera & K.A. Holen. 2017. A 130,000-year-old archaeological site in southern






https://doi.org/10.15184/aqy.2019.14 Published online by Cambridge University Press

California. *Nature* 544: 479–83. https://doi.org/10.1038/nature22065

– 2018a. Holen *et al.* reply. *Nature* 554: E3. https://doi.org/10.1038/nature25166

Holen, S.R., T.A. Deméré, D.C. Fisher, R. Fullagar, J.B. Paces, G.T. Jefferson, J.M. Beeton, R.A. Cerutti, A.N. Rountrey & K.A. Holen. 2018b. Broken bones and hammerstones at the Cerutti Mastodon site: a reply to Haynes. *PaleoAmerica* 4: 8–11. https://doi.org/10.1080/20555563.2017.1396835

Holen, S.R., T.A. Deméré, D.C. Fisher, R. Fullagar, J.B. Paces, G.T. Jefferson, J.M. Beeton, A.N. Rountrey & K.A. Holen 2018c. Disparate perspectives on evidence from the Cerutti Mastodon site: a reply to Braje *et al.* *PaleoAmerica* 4: 12–15. https://doi.org/10.1080/20555563.2017.1396836

Lesnek, A., J. Briner, C. Lindqvist, J. Baichtal & T. Heaton. 2018. Deglaciation of the Pacific coastal corridor directly preceded the human colonization of the Americas. *Science Advances* 4 (5): eaar5040.

Lin, S.C., Z. Rezek & H.L. Dibble. 2017. Experimental design and experimental inference in stone artifact archaeology. *Journal of Archaeological Method and Theory* 25: 633–88.

Magnani, M., Z. Rezek, S.C. Lin, A. Chan & H.L. Dibble. 2014. Flake variation in relation to the application of force. *Journal of Archaeological Science* 46: 37–49. https://doi.org/10.1016/j.jas.2014.02.029

Mesoudi, A. 2011. *Cultural evolution: how Darwinian theory can explain human culture and synthesize the social sciences*. Chicago (IL): University of Chicago Press. https://doi.org/10.7208/chicago/9780226520452.001.0001

Moreno-Mayar, J. *et al.* 2018. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* 553: 203–207. https://doi.org/10.1038/nature25173

Pedersen, M. *et al.* 2016. Postglacial viability and colonization in North America's ice-free corridor. *Nature* 537: 45–49. https://doi.org/10.1038/nature19085

Potter, B., J. Baichtal, A. Beaudoin, L. Fehren-Schmitz, C. Vance Haynes & V. Holliday. 2018. Current evidence allows multiple models for the peopling of the Americas. *Scientific Advances* 4(8): eaat5473.

Raghavan, M. *et al.* 2015. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* 349: aab3884. https://doi.org/10.1126/science.aab3884

Sankararaman, S., S. Mallick, N. Patterson & D. Reich. 2016. The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Current Biology* 26: 1241–47. https://doi.org/10.1016/j.cub.2016.03.037

Skoglund, P. & D. Reich. 2016. A genomic view of the peopling of the Americas. *Science* 41: 27–35.

Yravedra, J., J. Aramendi, M.A. Maté-González, L. Courtenay & D. González-Aguilera. 2018. Differentiating percussion pits and carnivore tooth pits using 3D reconstructions and geometric morphometrics. *PLoS ONE* 13: e0194324. https://doi.org/10.1371/journal.pone.0194324

*Received: 15 June 2018; Revised: 25 September 2018; Accepted: 22 October 2018*

Debate





https://doi.org/10.15184/aqy.2019.14 Published online by Cambridge University Press