

Another tool in the experimental toolbox: On the use of aluminum as a substitute for chert in North American prehistoric ballistics research and beyond

Metin I Eren, Lawrence Mukusha, Julie Lierenz, Michael Wilson, Michael Fisch, Trent True, Michael Kavaulic, Robert S. Walker, Briggs Buchanan, Alastair Key, Michelle R. Bebbler

Abstract:

Experimental archaeology continues to mature methodologically and theoretically. Around the world, practitioners are increasingly using modern materials that would have been unavailable to prehistoric people in archaeological experiments. The use of a modern material substitute can offer several benefits to experimental method, design, control, replicability, feasibility, and cost, but it should be directly compared to its “traditional” analogue to understand similarities and differences. Here, aluminum is introduced as a substitute for chert in prehistoric ballistics research because, critically, aluminum is safe, inexpensive, easy to process, and it and chert possess densities that differ by less than 4%. The aluminum casting process for replicating stone artifacts is presented, and it is shown that the aluminum castings are essentially identical in form, flake-scar patterning, and mass to their stone counterparts. We then present a proof-of-concept ballistics experiment that demonstrates no difference between aluminum and stone points in terms of target penetration.

1. Introduction

Experimental archaeology – archaeology's sub-discipline charged with recreating and reverse engineering ancient technologies to understand their manufacture and function – has a long, fruitful history in archaeological analysis and interpretation (e.g. Coles 1979; Johnson 1978; Lamdin-Whymark 2009; Meltzer 2015; O’Sullivan et al. 2014; Whittaker 1994, 2004). However, in the last two decades, experimental archaeology has substantially matured theoretically and methodologically (Eren and Bebbler 2019; Eren et al. 2016; Ferguson 2010; Jennings et al. 2021; Lin et al. 2018; Lycett and Chauhan 2010; Lycett and Eren 2013; Magnani et al. 2019a; Outram 2008; Rademakers et al. 2021; Schiffer 2013; Souyouzoglou-Haywood and O’Sullivan 2019). This maturity is manifested by several now-routine conventions: explicit hypothesis testing; better consideration of experimental design, variables, and variable interaction; an increase in experimental sample sizes or trials; the regular use of quantitative and statistical methods; multidisciplinary collaboration; situating experimental questions and results within an evolutionary framework; and publishing raw experimental data. There is still much work to do, but concomitant with these recent positive developments is the steady neutralizing of craftsperson authority, expertise, personal experience, and intuition that, for far too long, dominated experimental archaeology and suppressed the sub-discipline’s potential towards contributing to a robust understanding of technological adaptation and evolution¹ (Eren et al. 2016; Thomas 1986).

One increasingly common convention practiced by experimental archaeologists is the use of modern material substitutes for the types of materials that would have been more likely available to, or experienced by, past peoples. For example, knappers have long used copper pressure flakers or percussion “boppers” as a substitute for antler pressure flakers or billets (Whittaker 1994, 2004). Experimenters have used both glass and porcelain as substitutes for conchoidally fractured stone (Dibble and Pelcin 1995; Dibble and Rezek 2009; Dogandžić et al.

2020; Hecht et al. 2015; Iovita et al. 2014, 2016; Khreisheh et al. 2013; McPherron et al. 2020; Rezek et al. 2011; Speer 2018; Stout et al. 2015), and foam, plasticine, and even potatoes have been used as substitutes for stone specifically or reductive materials more generally (Clarkson 2017; Schillinger et al. 2014a, 2014b, 2015, 2016, 2017). Some recent experiments have featured conchoidally fracturing stone specimens that were ground into shape with modern lapidary equipment, rather than knapped (e.g. Eren et al. 2020, 2021; Lowe et al. 2019). Super glue and modern thermoplastic adhesive have been used as a substitute for pine pitch or hide glue in the hafting of stone tips to wooden shafts, and waxed twine has been used as a substitute for sinew or other types of lashing (Pargeter 2007; Wilson et al. 2021). Experiments assessing projectile penetration have used clay, foam, and silica gel targets as substitutes for live targets or carcasses (e.g. Key et al. 2018; Loendorf et al. 2018; Milks et al. 2016; Schoville et al. 2017; Wilkins et al. 2014) and wooden boards have been used in projectile durability tests in lieu of trees (e.g. Lowe et al. 2019; Eren et al. 2021a). Cutting experiments have employed the use of plastic tubing, modern ropes, cardboard, or bags of clay as cutting substrates (e.g. Bebbler et al. 2019; Key and Lycett 2011, 2014, 2015, 2017; Key et al. 2018). And we would be remiss in not mentioning equipment used in experimental archaeology like compound bows, projectile launchers, materials testers, electric kilns, wind tunnels, ball bearings, forges, etc, all of which are modern substitutes for past equipment that likely would have been unavailable and/or manufactured from different materials, in some cases vastly so (e.g. Bebbler 2017; Coppe et al. 2019; Dibble and Whittaker 1981; Lipo et al. 2012).

The use of modern material substitutes can provide several advantages to experimental design, interpretation, and replication. First, such substitutes can enhance the internal validity of an experiment and allow investigators to better isolate particular variables of interest (Eren et al. 2016; Lin and Premo 2021; Lin et al. 2018; Lycett and Eren 2013; Mesoudi 2011). Second, the use of modern material substitutes may assist in increasing experimental sample size and statistical robustness since these substitutes may be mass-produced or commercially available. Third, such mass-production or commercial availability can facilitate the replicability of an experiment by other researchers or laboratories. Fourth, modern material substitutes may be inexpensive to produce or use in terms of cost, skill, or time relative to those produced or used by past peoples. For example, producing a ‘natural’ adhesive (e.g. birch tar, Kozowyk et al. 2017; Schmidt et al. 2019) for hafting stone points onto projectile shafts may take skill or time, or introduce undesirable variability between trials or tests, that could be eliminated via the use of a modern adhesive. Finally, use of modern materials allows archaeologists to perform experiments that would otherwise be difficult or impossible. The use of porcelain cores, for instance, has allowed archaeologists to examine how different knappers flake cores of essentially the same form.

For all the advantages that modern material substitutes can offer to the practice of experimental archaeology, there are drawbacks too. Worth emphasizing is the fact that experimental control is a strategy in which any perceived benefit from one degree or kind of control necessarily comes with an unavoidable cost (Eren et al. 2016:106-107). Some experiments may require, or gain from, an analysis of “noise”, while others may be interested in analyzing the interaction of particular variables (compare Lipo et al. 2012 and Pettigew et al. 2015). Furthermore, the perceived control and uniformity that can come with the use of modern materials may introduce or deactivate unknown or hidden variable interactions that can make it difficult to directly apply experimental findings to the interpretation of the archaeological record. For example, using a vice, rather than a human hand, to hold a stone core for stone flake removal

introduces a variable that was not present in the past (i.e. the vice) that may remove or alter the human interaction archaeologists are ultimately interested in, even though important information can still be obtained. In other words, inherent to the use of modern material substitutes in archaeological experiments is an automatic and unavoidable decrease in external validity, a decreased “realism”, that requires some sort of theoretical or practical justification, or an explicit proviso that acknowledges the indirect applicability of results to the past relative to experiments that use more “traditional” materials (i.e. decreased “realism”, Eren et al. 2016; Lycett and Eren 2013; Mesoudi 2011; Pettigrew et al. 2015). In sum, arguments for “realistic” experiments being better than “controlled” ones, or *vice versa*, are in vain. Perhaps a “controlled” experiment can answer a particular question better than a “realistic” one – and again, *visa versa* – but that is different than claiming control or realism is always better. Realistic and controlled experiments (and everything in between) provide different types of information, can answer different types of questions, and results from each kind of experiment should be juxtaposed against the other. A robust experimental archaeology needs the entire external-internal validity spectrum (Eren et al. 2016).

The direct comparison of modern material substitutes to their traditional analogues can go a long way toward justifying the former’s use in archaeological experiments. Even if a modern substitute does not behave entirely the same as a traditional material – indeed, they rarely do – understanding exactly how the two materials are similar and different can help bolster confidence in experimental results and their subsequent applicability to the archaeological record. Or, understanding how modern and traditional materials are similar and different can help determine in what capacity or for which particular experiment a modern substitute can be reasonably employed. Toward this end, several studies in recent years have directly compared modern materials to traditional ones to gain such an understanding² (e.g. Dogandžić et al. 2020; Key et al. 2018; Lowe et al. 2019; Wilson et al. 2021).

Here we introduce the use of aluminum as a modern material substitute for chert in prehistoric ballistics research, and present an experiment that validates aluminum’s use in some experimental contexts. Aluminum can be cast into precise and accurate replicas of stone points, and, importantly, *aluminum (~2.7 g/cm³) and chert (~2.6 g/cm³) are nearly identical in density*. Thus, aluminum and chert projectile points, all else being equal, should in some ways perform similarly in terms of ballistics, for example velocity and penetration, which is our experimental focus in this study. Aluminum is also easier and safer to handle, and less expensive, than other alloys of density similar to chert³.

2. Materials and methods

2.1 Knapping bifacial stone points

M.I.E. knapped three chert points using percussion and pressure flaking. The first point was a fluted Clovis point knapped from Texas Georgetown chert (**FIGURE 1a**). The second point was a triangular Madison point knapped from Ohio Upper Mercer chert (**FIGURE 1b**). The third point was a Side-Notched point knapped from French Bergerac flint (**FIGURE 1c**).

2.2 Production of aluminum casts

We replicated the chert points – producing three casts each of the three point types – using the loose pattern casting process in the foundry at the College of Aeronautics and Engineering at Kent State University. We used an Oil Bonded molding sand called K-Bond that was developed at Kent State in the spring of 1992 to replace commercially available oil bound sands that used recycled motor oil. The first step in working with this sand is to recondition it from its last use. We shoveled the sand from a molding bench to a Carver sand muller to be reconditioned and remixed. We added an oil swelling clay (**FIGURE 2a**) and synthetic oil (Amorphous Polybutene) (**FIGURE 2b**) and allowed the batch to mix in the muller until it was homogeneous. We then transferred the sand to a hopper (**FIGURE 2c**) and took it back to the molding bench to be used.

2.2.1 Molding

Tools used in the molding process include backing boards, cope (top) and drag (bottom) molding flasks (containers with four sides), dry parting compound, a riddle (sieve), hand ram for compacting the sand, strike-off bar, sprue cutter (tool for making a hole to pour liquid metal into the mold), and pouring basin mold (basin above the sprue that is a reservoir of liquid metal during pouring) (**FIGURE 3**).

The first step is to place a backing board on the molding bench, and then placing the drag half of the molding flask on the board. The chert points to be reproduced are then placed within the extents of the flask, making sure to think about the placement of the gating system (the path the liquid metal will follow from outside to inside the mold) (**FIGURE 4a**). Once everything was set, the points and backing board were dusted with dry parting compound. Parting compound stops the molding sand from sticking to the pattern and backing boards (**FIGURE 4b**).

Sand was screened through a riddle to remove any debris or metal scraps from its last usage. This ensured there was nothing but molding sand in contact with the surface of the stone point (**FIGURE 4c**). This initial layer of sand, referred to as ‘facing sand,’ was then gently hand compacted with care to not move or disturb the points (**FIGURE 4d**). Additional sand was screened through the riddle, added to the flask and compacted using a hand ram (**FIGURE 5a**). The sand added in this step is referred to as backing sand as it does not touch the pattern; this process was repeated until the compacted sand was slightly higher than the edges of the flask half. Using a strike-off bar the sand level was then scraped to the height of the flask (**FIGURE 5b**). A second backing board was placed on top of the drag mold and then the flask half and both backing boards flipped over. The current top backing board is then removed thus exposing the points (**FIGURE 5c**).

Due to the curved and irregular shape of the stone points it was necessary to use a technique called offset molding to duplicate them. In this process one identifies the edges of the stone point and removes any sand that had wrapped around the pattern during the first step described above (**FIGURE 5d**). A cut at a shallow angle from the surface of the sand to the edges of the pattern is then made (**FIGURE 6a**). A second coating of parting compound was then applied to stop the sand from sticking to itself and the back side of the pattern (**FIGURE 6b**). The cope half of the molding flask is then placed on the drag half and the molding process of adding sand and compressing is repeated. After the cope has been struck off, we began the creation of the gating system by cutting the sprue hole (a tapered hole through the cope half of the mold terminating slightly in the drag) so when the mold is opened, we could easily locate

where the sprue aligns (**FIGURE 6c**). The mold was then opened, points removed, and the gating system is extended from the sprue location to the points. The drag area around the sprue hole was opened and made deeper to create a sprue basin. This open space helps to lower the velocity of the metal generated by dropping down the sprue (**FIGURE 6d**).

Using compressed air, we blew any loose molding sand off the mold halves before closing them. A pouring basin is added to the top of the mold and fixed in place by surrounding it with molding sand and hand compacting to cause the sand to stick to both the top of the mold and the side walls of the pouring basin (**FIGURE 7a**).

The completed molds are transferred to the pouring floor (**FIGURE 7b**) and clamped to keep the mold halves from separating during pouring.

2.2.2 Melting & Pouring

We began by loading the furnace with smaller pieces of scrap/remelt aluminum. All the aluminum we used is grade A356.1 alloy (**FIGURE 7c**). Once a liquid puddle formed in the bottom of the furnace additional melt stock was added until the desired volume of metal was liquid in the furnace (**FIGURE 7d**). Temperature of the liquid bath is measured using an immersion pyrometer to ensure it was at an appropriate temperature for pouring (**FIGURE 8a**). Metal was then transferred into a pouring ladle and then poured into the molds (**FIGURE 8b & 8c**).

2.2.3 Shake Out & Clean Up

After allowing the molds to cool and the metal to solidify, the parts are removed from the molds and the sand returned to the casting benches to be recycled to make new molds (**FIGURE 8d**). The cooled molds are moved back to the casting benches, opened the sand broken up and the cast points and attached gating system removed (**FIGURE 9a**). The cooled castings were taken to a tub to remove more sand (**FIGURE 9b**), then to a band saw for removal of the gating system (**FIGURE 9c**), and then shaped using a belt sander to hide the attachment point for the gating system. A Dremel tool was used to remove extra aluminum imperfections (**FIGURE 9d**).

2.3 Comparison of aluminum and chert density

. Given the relationship between mass (M), volume (V), and density (d), which is ($M=Vd$), it is important to reiterate that the reason aluminum is a potential substitute for chert in some prehistoric ballistics research because the density of the two materials' density is similar. To confirm this, we compared the density of our grade A356.1 aluminum alloy with Georgetown, Bergerac, and Upper Mercer cherts using an Alfa MirageTM MD 300S densimeter. The measurements were carried out at 19 and 20 °C and the temperature and calibration verified before measurements. The density was corrected to 15 °C. Four to six measurements of density were made on each aluminum sample and three measurements on each flint samples. The mean density of the aluminum was 2.65 g/cm³ and was essentially the same for all the points. The mean density of Georgetown chert was 2.55 g/cm³; of Bergerac chert was 2.56 g/cm³; and of Upper Mercer chert was 2.56 g/cm⁴.

2.4 Hafting of stone and aluminum points

We randomly selected one Clovis, Madison, and Side-Notched aluminum point to be hafted in the same manner as its knapped stone point counterpart (**FIGURE 10**) (and it was this single hafted aluminum specimen of each point type that was shot). We used the same basic procedure to haft all six specimens, but the aluminum and stone Clovis points required a larger diameter shaft, while the aluminum and stone Madison and Side-Notched points were hafted on smaller diameter shafts. We first rough-cut a pocket into each shaft using a carpenter's coping saw. We then fine-tuned and pre-fitted the haft via sharp blades and files. Next, the pocket and leading end of the shaft were heat-hardened with the flame of Sterno™ canned heat gel. The basal ends of the aluminum and stone point were warmed while heating a thermoplastic adhesive to its melting point (Wilson et al. 2021). After fully wetting the mating surfaces (the pocket and the basal end of the projectile), we inserted the point into the shaft pocket, centered it while cooling, and then spin-balanced the projectile by hand. We filled any voids in the pocket area, resulting in a streamlined convexity on each of the projectile's faces.

After the glue was set, binding was applied. For binding the haft sites, we chose an artificial sinew. This product is a multi-strand polyester filament, coated in beeswax and rated at 70lb test strength. Although this product does not shrink, it works very well to prevent shafts splitting upon impact. However, we note that experimental comparisons should be conducted comparing this artificial sinew to other types of "natural" lashing. A basic whip knot (used in naval rope rigging) served as the binding lashing and knot. We then warmed the entire haft site again for final smoothing and fine-centering. Finally, we applied three coats of water-based polyurethane sealer to the haft site to seal and protect the binding.

We encountered a unique challenge to the hafting process of aluminum points that we had not anticipated (but perhaps should have). Although aluminum and chert share similar densities, and the former can be cast to appear nearly identical in form and flake-scar patterning to the latter, we failed to foresee that aluminum is an excellent conductor of heat, while the stone is not. Thus, while the chert stone heated gradually and maintained the heat longer (making it easier to handle), the aluminum became very hot, very fast, making it extremely challenging to handle while hafting. While obvious in hindsight, we report this noticeable (and painful) difference between aluminum and stone temperature here so that other researchers can be prepared with proper equipment as needed (e.g. heat-proof gloves).

2.5 Shooting stone and aluminum points

We shot the hafted aluminum and stone points in the Kent State University Experimental Archaeology Laboratory following previously published equipment and procedures (see Eren et al. 2020, 2021b; Mullen et al. 2021; Sitton et al. 2020 for detailed description and pictures). We used a Gamma Master Model Shooting Chronograph to measure velocity of each projectile in meters per second (m/s) throughout the experiment (Eren et al. 2021a). The chronograph readings result in "error" if there is a change in sunlight, cloud cover, or some other minor variable. As a result, we recorded a percentage of 30 possible stone point velocity readings per aluminum or stone point type. As predicted, the velocities of the aluminum and stone points in each set were nearly identical given that the projectile masses were similar and each projectile was pulled to the same draw length (**TABLE 1**).

We recorded penetration depth into the clay target for each shot and measured it by marking the shaft at the location at which the shaft was exposed (Key et al. 2018). Once the

projectile was removed, we used a tape measure to measure in centimeters (cm) from the mark on the shaft to the tip of the projectile point. Samples sizes for each projectile are shown in **TABLE 1**.

2.6 Statistical analysis

We ran Bayesian equivalents of the two sample t-test (Kruskale 2013) to test penetration depth differences between aluminum and stone points for each our point types of Clovis, Side-Notched, and Madison. Our distributional t-tests allow unequal variance between groups and are run in *R* 4.0.5 (*R* Core Team) with the *brms* package. Following Kruskale (2013), we used a weak exponential prior with a rate of 1/29 for *nu*, the parameter that changes the normality of the distribution (i.e., the degrees of freedom parameter in the *t* distribution) and a cauchy (0,1) prior for the variance, or sigma, in each group. Final models were run with 2 chains for 10,000 iterations after a warmup of 5,000 iterations. Rhat values (a model diagnostic with expected value equal to 1) for all parameters were exactly 1.00 and hence ensure model convergence. Chains were also inspected visually for sufficient mixing to ensure that model results were appropriate.

All raw data and *R* scripts are available in the supplementary online materials.

3. Results

3.1 Stone versus aluminum cast form, mass, and flake scar patterning

TABLE 2 presents a comparison of basic measurements between the aluminum casts and knapped stone point models. **FIGURE 11** shows the aluminum and stone points side-by-side. These measurements and image demonstrate that precise and accurate stone point replicas – in terms of form, mass, and flake scar patterning – can be produced from aluminum casting. While these results are positive, we anticipate even better aluminum casts will be made as we refine the process further.

3.2 Comparison of penetration depth

TABLE 1 presents the mean penetration depths for the aluminum and stone Clovis, Madison, and Side-Notched projectile point sets. We carried out three separate Bayesian analyses of the penetration depths comparing Clovis, Madison, and Side-Notched points made of aluminum to points made of stone. The results indicated that Clovis and Madison points had similar penetration depths, while the stone Side-Notched points penetrated deeper than aluminum Side-Notched points (**FIGURE 12; TABLE 3**; for full result tables see Supplementary Materials).

Because we expected that all three point types would have statistically similar penetration depths we subsequently reshot the Side-Notched aluminum and stone points in a second trial. The results of this work also showed that the stone and aluminum Side-Notched points had statistically different penetration depths (**FIGURE 13; TABLE 3**).

We were puzzled by this result because the point form, flake-scar patterning, projectile mass, and velocity were nearly identical. Upon close examination of the two Side-Notched projectiles, we noticed a small, but observable difference in the bindings (**FIGURE 14**). The

aluminum projectile's bindings were rough on each face from an uneven application of polyurethane sealer, whereas the stone projectile's bindings were smooth. We hypothesized that *because* the point form, flake scar patterning, projectile mass, and velocity were identical, this difference in hafting was causing the significant difference in penetration depth. In other words, the aluminum and stone projectiles were so similar, we were able to unexpectedly identify a minor, but significant contributor to penetration depth: lashing roughness. To test this hypothesis, we re-hafted the aluminum and stone Side-Notched points without lashings, only using adhesive. In a third round of shooting, we shot each Side-Notched projectile 30 times. Finally, as predicted, our Bayesian model showed no significant penetration depth difference (**FIGURE 13; TABLE 3**).

4. Discussion

As archaeological fieldwork declines in future years for ethical, financial, or practical reasons (e.g. Kletter and de Groot 2001; Nash and Colwell 2020; Surovell et al. 2017) archaeologists can continue to look forward to better understanding the past by focusing their efforts on the vast, understudied, and already-excavated collections amassed over the last few centuries. One way to productively and cost-effectively approach the study of these collections is via experimental archaeology (Diez-Martin et al. 2021; Magnani et al. 2019b), which has recently matured and expanded at a furious pace. An increasingly fruitful and highly replicable practice in experimental archaeology is the use of proxy materials: modern material substitutes for the types of materials that would have been more likely available to, or experienced by, past peoples. However, before their use in archaeological experiments, modern material substitutes should be directly compared with more “traditional” materials to understand in what particular ways they are similar and different. Here we introduced the use of aluminum as a substitute for stone given that the two materials share nearly identical densities, and the former can be poured into molds created from the latter resulting in duplicate form, flake scar patterning, and mass. As proof-of-concept, we shot three sets of projectiles (Clovis, Madison, Side-Notched), each made from aluminum and stone, and compared them ballistically. Our results showed that for ballistics tests involving velocity and penetration aluminum can be used as an accurate, precise, and reliable substitute for chert.

It is worth emphasizing that while aluminum can serve as a substitute for chert in some types of experimental ballistics tests, it would not do well in others. On one hand, experimental assessment of stone point durability or sharpness, for example, might be poorly served by the use of aluminum points. However, other tests beyond penetration – such as aerodynamics, flight trajectory, or other applications we have yet to think of – might profit from the use of aluminum points, especially since aluminum points will not likely break or incur distal impact scars. Indeed, we encountered an accidental stone point break during the experiments we reported here. One of us (M.I.E.) dropped the stone Madison point on its 28th shot, and hence the sample size for that comparison was smaller than the target sample of 30 (**TABLE 2**).

There are other potential benefits to the use of aluminum points. First, real, prehistoric stone artifacts can be cast in aluminum and used in appropriate ballistics tests. For example, rather than using a flintknapper to knap a stone point for an experimental ballistics test, an actual Clovis point from Blackwater Draw (New Mexico), or a Mississippian point from Cahokia (Illinois) could be cast in aluminum and shot⁵. In other words, casting points in aluminum essentially allows archaeologists to shoot genuine “stone” artifacts in *some* experimental tests.

Questions or skepticism about whether the flintknapper is accurately replicating artifacts can be eliminated because, for all intents and purposes, the “real” artifact is being tested. The use of aluminum, therefore, can increase the experimental external validity of some tests in particular ways when appropriate or necessary while simultaneously increasing internal validity in other specific ways (Eren et al. 2016; Lin et al. 2018; Lycett and Eren 2013; Mesoudi 2011).

Researchers may not have permission or immediate access to cast genuine stone artifacts, and thus a second benefit to the use of aluminum as a substitute for chert is the ability to turn a plastic cast or resin 3D print “back into stone.” Casts are widely available for purchase or loan and inexpensive to ship, while 3D scanning and printing is now readily available to most. While more research into the plastic-to-aluminum cast process is necessary, we successfully conducted a preliminary test using a Clovis point cast from the Lamb site (specimen 86/94/106, New York, U.S.A.) (**FIGURE 11**).

A third potential benefit to the use of aluminum as a substitute for stone is for studying real or hypothetical scenarios involving the perishable aspects of prehistoric weapon systems: shafts, lashings, fletching, and adhesives (Eren et al. 2021a; Whittaker 2010). For instance, several aluminum copies of a single stone artifact can be created for the purposes of understanding which type of hafting configuration is the most durable or will best facilitate penetration. Indeed, we were only able to tease out the unexpected and significant effect of lashing roughness on target penetration in our experiments (Section 3.3) because the aluminum and stone points were nearly identical in their form, flake scar patterning, and mass.

The advantages of using aluminum casts are not limited to lithic projectile experiments. Investigations into form-related attributes (e.g., performance) of hand-held technologies including flakes, handaxes, blades, and scrapers could similarly benefit. Indeed, “real” Acheulean handaxes, Mousterian scrapers, or Upper Palaeolithic burins, for example, could be used after being cast using artifact molds or 3D printouts. Not only would this allow individuals to use the precise artifact forms created by past individuals – the benefits of which have already been discussed – but it would allow the tool designs imposed by extinct hominin species to be used in some experimental tests, albeit related to a limited number of topics. For the first time, it would be possible for tools designed and produced by species with cognition and anatomy distinct to our own to be used. As it stands, experimental lithic archaeology relies on best guess-solutions to past tool-design criteria and production techniques as inferred by modern humans. Often replica lithic assemblages are a good solution covering many forms seen in the artifact record (Eren et al. 2016, 2020, 2021), but as recently displayed for Acheulean handaxes, even substantial replica tool assemblages (e.g., >500) specifically designed to meet and then go beyond tool-forms observed in the artifact record can still miss design elements imposed by extinct species (Key 2019). The close matching density of aluminum and chert, and the precise nature of cast replications, would have particular advantages during ergonomic investigations into artifact design and use. Indeed, finer considerations of tool prehension could be accurately replicated. This includes the friction provided by flake scars, the added ‘comfort’ that specific forms of ‘backing’ provides when gripped, and the tool’s overall balance and mass distribution. All of which are essential to accurately replicating the biomechanical stresses when tools are wielded.

Some readers may wonder how widely available the facilities or materials are for aluminum casting. Aluminum is relatively inexpensive, and in North America Foundries or Casting Facilities are common. There are a dozen or so such facilities in the Cleveland/Akron (Ohio) metropolitan area alone, and while this frequency likely varies geographically, we suggest

that it may not be difficult to find appropriate facilities, or to work with collaborators who can. Of course, flintknappers may not be difficult to find either in certainly countries like the USA that possess strong hobby-knapping communities. In other countries finding knappers may be more difficult. Time-wise, making aluminum casts may be quicker or slower than stone point knapping depending on the point type. It took an afternoon to make our 16 aluminum casts. Making 16 small Madison points in that time may be possible, although they likely would not be highly standardized. Yet, making sixteen Clovis points might take a two or more work days. Cost-wise, aluminum and chert are both relatively inexpensive materials – 50 lbs. of aluminum costs \$257.00 (<https://www.midwesttechnology.com/kroger-aluminum-alloy-a356-ingot-50-lbs/>, accessed 4 November 2021) and 50 lbs. of Texas chert costs \$137.50 (<https://www.neolithics.com/raw-rock-based-on-pounds-not-by-piece/>, accessed 4 November 2021).

Given these various considerations, one may also wonder about the costs and benefits of using aluminum versus stone in an experiment. We suggest that this ultimately depends on the question being tested. If the experiment at hand will best be served using an aluminum cast of an actual artifact that will not likely snap in half or break, then aluminum should be used. Similarly, if the experiment at hand will best be served by a stone replica made by a modern knapper, then stone should be used. When we suggest that aluminum be considered as a substitute for stone, we do not mean in every testable instance – only in experimental tests where it might be beneficial in some way. Hence, aluminum should be seen as another tool in the experimental archaeologist's toolbox – not the only tool, and not an automatically superior tool. We look forward to future tests and new applications to see how this new tool can contribute to our understanding of lithic technology, prehistoric ballistics, and cultural evolution.

Endnotes

1. This is not to say that a modern craftperson's expertise, experience, or intuition is unimportant or without value, only that expertise, experience, and intuition cannot be the sole basis for a robust, scientific understanding of the past. Expertise, experience, and intuition are excellent tools for developing hypotheses, but cannot serve as the test or proof of those hypotheses.

2. A popular modern material substitute for conchoidally-fracturing rock is porcelain (Hecht et al. 2015; Khreisheh et al. 2013; Speer 2018; Stout et al. 2015; Tsirk 2014), which is knapped or broken in specific ways for various experimental archaeological investigations. While we fully acknowledge the advantage and insights that porcelain can provide to archaeological experiments, we are surprised that – to our knowledge – there has not been a comparison of amount of force necessary to initiate a crack in porcelain versus stone. Given that the use of porcelain is often used in knapper skill-level studies, it should be considered that if the amount of force to initiate a crack in porcelain is different than it is in chert (which is what we currently suspect), knapper skill level may possibly be inflated, deflated, or in some way confounded since hammer speed may differ, and thus so too may hand-eye-coordination (Mraz et al. 2018; Stout 2005). Or, a porcelain cast of a bifacial point may more easily break relative to a stone analogue, which may skew interpretations of artifact function if only the former are experimentally tested (e.g. Khreisheh et al. 2013). The key is to make sure that for any particular

test the modern material substitute is suited to the performance task, or at least any differences between the substitute and the “archaeological” materials are understood. In other words, porcelain’s possession of conchoidal fracture does not automatically mean it “behaves” entirely like stone. The same can be said of glass (e.g. Iovita et al. 2014), which, while possessing conchoidal fracture, requires at a smaller force to initiate a crack than does stone (Dogandžić et al. 2020).

3. For the interested reader, the Young’s Elastic Modulus of A356.1 aluminum alloy and unheated French chert is 73-80 GPa and 75-76, respectively (<https://matmatch.com/materials/minfm12662-astm-a356-0-t6>, accessed 4 November 2021; Schmidt et al. 2019).

4. Aluminum shrinks by over 6% during solidification. Thus, a casting is dimensionally smaller (in all three perpendicular directions) by roughly 2% than the mold. Because of the sand grains finite size and surface tension between the sand and liquid metal. A typical surface finish is ± 0.007 to 0.015 mm, and a dimension tolerance is $\pm .76$ mm.

5. Or, expanding beyond North America to the Old World, an actual Levallois point from Qafzeh Cave (Israel) could be cast in aluminum and shot (Iovita et al. 2014).

References

Bebber, M. R., 2017. Tempered Strength: a controlled experiment assessing opportunity costs of adding temper to clay. *Journal of Archaeological Science*, 86, 1-13.

Bebber, M. R., Key, A. J., Fisch, M., Meindl, R. S., Eren, M. I. 2019a. The exceptional abandonment of metal tools by North American hunter-gatherers, 3000 BP. *Scientific reports*, 9(1), 1-4.

Clarkson, C., 2017. Teaching Complex Flint Knapping Strategies in the Classroom Using “Potato Knapping”. *Lithic Technology*, 42, 155-160.

Coles, J., 1979. *Experimental Archaeology*. Academic Press, London.

Coppe, J., Lepers, C., Clarenne, V., Delaunois, E., Pirlot, M., & Rots, V., 2019. Ballistic Study Tackles Kinetic Energy Values of Palaeolithic Weaponry. *Archaeometry*, 61, 933-956

Dibble, H. L., & Pelcin, A., 1995. The effect of hammer mass and velocity on flake mass. *Journal of Archaeological Science*, 22, 429-439.

Dibble, H. L., Rezek, Z., 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, 1945-1954.

Dibble, H. L., Whittaker, J. C., 1981. New experimental evidence on the relation between percussion flaking and flake variation. *Journal of Archaeological Science*, 8, 283-296.

Diez-Martin, F., Buchanan, B., Norris, J. D., Eren, M. I. (2021). Was Welling, Ohio (33-Co-2), a Clovis Basecamp or Lithic Workshop? Employing Experimental Models to Interpret Old Collections. *American Antiquity*, 86(1), 183-198.

Dogandžić, T., Abdolazadeh, A., Leader, G., Li, L., McPherron, S. P., Tennie, C., Dibble, H. L., 2020. The results of lithic experiments performed on glass cores are applicable to other raw materials. *Archaeological and Anthropological Sciences*, 12, 1-14.

Eren, M. I., Lycett, S. J., Patten, R. J., Buchanan, B., Pargeter, J., O'Brien, M. J., 2016. Test, model, and method validation: the role of experimental stone artifact replication in hypothesis-driven archaeology. *Ethnoarchaeology*, 8, 103-136.

Eren, M. I., Bebber, M. R. 2019. The Cerutti Mastodon site and experimental archaeology's quiet coming of age. *Antiquity*, 93, 796-797.

Eren, M. I., Story, B., Perrone, A., Bebber, M., Hamilton, M., Walker, R., Buchanan, B. (2020). North American Clovis point form and performance: An experimental assessment of penetration depth. *Lithic Technology*, 45(4), 263-282.

Eren, M. I., Romans, J., Buchanan, B., Bebber, M. R. 2021. Validating chronograph photo sensor measurement accuracy of stone-tipped projectile velocity. *Measurement: Sensors*, 13, 100037.

Eren, M. I., Bebber, M. R., Wilcox, D., Story, B., Buchanan, B., 2021a. North American Clovis Point Form and Performance II: An Experimental Assessment of Point, Haft, and Shaft Durability. *Lithic Technology*, In Press.

Eren, M. I., Meltzer, D. J., Story, B., Buchanan, B., Yeager, D., & Bebber, M. R. 2021b. On the efficacy of Clovis fluted points for hunting proboscideans. *Journal of Archaeological Science: Reports*, 39, 103166.

Ferguson, J. R. (Ed.), 2010. *Designing Experimental Research in Archaeology: Examining Technology Through Production and Use*. University Press of Colorado, Boulder.

Hecht, E. E., Gutman, D. A., Khreisheh, N., Taylor, S. V., Kilner, J., Faisal, A. A., Bradley, B., Chaminade, T., Stout, D., 2015. Acquisition of Paleolithic toolmaking abilities involves structural remodeling to inferior frontoparietal regions. *Brain Structure and Function*, 220, 2315-2331.

Iovita, R., Schönekeß, H., Gaudzinski-Windheuser, S., & Jäger, F. Projectile impact fractures and launching mechanisms: results of a controlled ballistic experiment using replica Levallois points. *Journal of Archaeological Science*, 48, 73-83.

Iovita, R., Schönekeß, H., Gaudzinski-Windheuser, S., & Jäger, F., 2016. Identifying weapon delivery systems using macrofracture analysis and fracture propagation velocity: a controlled

experiment, in: Iovita, R., Sano, K. (Eds.), *Multidisciplinary Approaches to the Study of Stone Age Weaponry*. Springer Netherlands, Dordrecht, pp. 13-27.

Jennings, T. A., Smallwood, A. M., Pevny, C. D., 2021. Reviewing the role of experimentation in reconstructing paleoamerican lithic technologies. *PaleoAmerica*, 7, 53-67.

Johnson, L. L., 1978. A history of flint-knapping experimentation, 1838-1976. *Current Anthropology*, 19, 337-372.

Key, A. J., Lycett, S. J. 2011. Technology based evolution? A biometric test of the effects of handsize versus tool form on efficiency in an experimental cutting task. *Journal of Archaeological Science*, 38, 1663-1670.

Key, A. J., Lycett, S. J., 2014. Are bigger flakes always better? An experimental assessment of flake size variation on cutting efficiency and loading. *Journal of Archaeological Science*, 41, 140-146.

Key, A. J., Lycett, S. J., 2015. Edge angle as a variably influential factor in flake cutting efficiency: an experimental investigation of its relationship with tool size and loading. *Archaeometry*, 57, 911-927.

Key, A. J., Lycett, S. J., 2017. Reassessing the production of handaxes versus flakes from a functional perspective. *Archaeological and Anthropological Sciences*, 9, 737-753.

Key, A., Young, J., Fisch, M. R., Chaney, M. E., Kramer, A., Eren, M. I., 2018. Comparing the use of meat and clay during cutting and projectile research. *Engineering Fracture Mechanics*, 192, 163-175.

Key, A. J. M., 2019. Handaxe shape variation in a relative context. *Comptes Rendus Palevol*, 18, 555-567

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, 38-46.

Kletter, R., De-Groot, A. 2001. Excavating to Excess?: The Last Decade of Archaeology in Israel and its Implications. *Journal of Mediterranean Archaeology*, 14, 76-85.

Kozowyk, P. R., Soressi, M., Pomstra, D., & Langejans, G. H. 2017. Experimental methods for the Palaeolithic dry distillation of birch bark: implications for the origin and development of Neandertal adhesive technology. *Scientific Reports*, 7(1), 1-9.

Kruschke, J., 2013. Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, 142, 573-603.

Lamdin-Whymark, H., 2009. Sir John Evans: experimental flint knapping and the origins of lithic research. *Lithics*, 30, 45-52.

- Lin, S. C., Rezek, Z., Dibble, H. L., 2018. Experimental design and experimental inference in stone artifact archaeology. *Journal of Archaeological Method and Theory*, 25, 663-688.
- Lipo, C. P., Dunnell, R. C., O'Brien, M. J., Harper, V., Dudgeon, J., 2012. Beveled projectile points and ballistics technology. *American Antiquity*, 77, 774-788.
- Loendorf, C., Blikre, L., Bryce, W. D., Oliver, T. J., Denoyer, A., Wermers, G., 2018. Raw material impact strength and flaked stone projectile point performance. *Journal of Archaeological Science*, 90, 50-61.
- Lowe, C., Kramer, A., Wilson, M., Meindl, R., Spurlock, L., Eren, M. I., 2019. Controlled ballistics tests of ground, percussion-flaked, and pressure-flaked projectile point impact durability: Implications for archaeological method and theory. *Journal of Archaeological Science: Reports*, 24, 677-682.
- Lycett, S. J., Chauhan, P. R., 2010. Analytical approaches to Palaeolithic technologies: an introduction. In: Lycett, S. J., Chauhan, P. R. (Eds.), *New Perspectives on Old Stones*, Springer, New York, pp. 1-22.
- Lycett, S. J., Eren, M. I., 2013. Levallois lessons: the challenge of integrating mathematical models, quantitative experiments and the archaeological record. *World Archaeology*, 45(4), 519-538.
- Magnani, M., Grindle, D., Loomis, S., Kim, A. M., Egbers, V., Clindaniel, J., Hartford, A., Johnson, E., Weber, S., Campbell, W., 2019a. Evaluating claims for an early peopling of the Americas: experimental design and the Cerutti Mastodon site. *Antiquity*, 93, 789-795.
- Magnani, M., Grindle, D., Loomis, S., Kim, A. M., Egbers, V., Clindaniel, J., Hartford, A., Johnson, E., Weber, S., Campbell, W. 2019b. Experimental futures in archaeology. *Antiquity*, 93, 808-810.
- McPherron, S. P., Abdolazadeh, A., Archer, W., Chan, A., Djakovic, I., Dogandžić, T., Leader, G., Li, L., Lin, S., Magnani, M., Reeves, J., Rezek, Z., Weiss, M., 2020. Introducing platform surface interior angle (PSIA) and its role in flake formation, size and shape. *PLoS ONE*, 15, e0241714.
- Meltzer, D. J., 2015. *The Great Paleolithic War*. University of Chicago Press, Chicago.
- Mesoudi, A., 2011. *Cultural Evolution*. University of Chicago Press, Chicago.
- Milks, A., Champion, S., Cowper, E., Pope, M., Carr, D., 2016. Early spears as thrusting weapons: Isolating force and impact velocities in human performance trials. *Journal of Archaeological Science: Reports*, 10, 191-203.

Mraz, V., Fisch, M., Eren, M. I., Lovejoy, C. O., Buchanan, B., 2019. Thermal engineering of stone increased prehistoric toolmaking skill. *Scientific Reports*, 9, 1-8.

Mullen, D., Matney, T., Morrison, A., Fisch, M., Buchanan, B., Bebbler, M. R., 2021. Experimental assessment of Neo-Assyrian bronze arrowhead penetration: An initial study comparing bilobate versus trilobate morphologies. *Journal of Archaeological Science: Reports*, 35, 102765.

Nash, S. E., Colwell, C. 2020. NAGPRA at 30: the effects of repatriation. *Annual Review of Anthropology*, 49, 225-239.

O'Sullivan A., Powers, M., Murphy, J., Inwood, N., Gilhooly, B., Kelly, N., Malone, W., Mulrooney, J., Corrigan, C., L'Estrange, M., Burke, A., Kazuro, M., McDermott, C., Warren, G., O'Neill, B., Heffernan, M., Sweeney, M., 2014. In: Kelly, B., Roycroft, N., Stanley, M. (Eds.), *Fragments of Lives Past: Archaeological Objects from Irish Road Schemes*, Archaeology and the National Roads Authority Monograph Series No. 11, pp. 115-126.

Outram, A. K., 2008. Introduction to experimental archaeology. *World Archaeology*, 40, 1-6.

Pargeter, J. 2007. Howiesons Poort segments as hunting weapons: experiments with replicated projectiles. *The South African Archaeological Bulletin*, 62, 147-153.

Pettigrew, D. B., Whittaker, J. C., Garnett, J., Hashman, P. 2015. How atlatl darts behave: beveled points and the relevance of controlled experiments. *American Antiquity*, 80, 590-601.

Rademakers, F., Verly, G., Téreygeol, F., Auenmüller, J. 2021. Contributions of experimental archaeology to excavation and material studies. *Journal of Archaeological Science: Reports*, 38, 103036.

Rezek, Z., Lin, S., Iovita, R., Dibble, H. L. 2011. The relative effects of core surface morphology on flake shape and other attributes. *Journal of Archaeological Science*, 38, 1346-1359.

Schiffer, M. B., 2013. Contributions of Experimental Archaeology. In: Schiffer, M. B., *The Archaeology of Science*, Springer, Heidelberg, pp. 43-52.

Schillinger, K., Mesoudi, A., Lycett, S. J., 2014a. Copying error and the cultural evolution of “additive” vs. “reductive” material traditions: an experimental assessment. *American Antiquity*, 79, 128-143.

Schillinger, K., Mesoudi, A., Lycett, S. J. 2014b. Considering the role of time budgets on copy-error rates in material culture traditions: an experimental assessment. *PLoS ONE*, 9, e97157.

Schillinger, K., Mesoudi, A., Lycett, S. J., 2015. The impact of imitative versus emulative learning mechanisms on artifactual variation: implications for the evolution of material culture. *Evolution and Human Behavior*, 36(6), 446-455.

- Schillinger, K., Mesoudi, A., Lycett, S. J. 2016. Copying error, evolution, and phylogenetic signal in artifactual traditions: An experimental approach using “model artifacts”. *Journal of Archaeological Science*, 70, 23-34.
- Schillinger, K., Mesoudi, A., Lycett, S. J. 2017. Differences in manufacturing traditions and assemblage-level patterns: the origins of cultural differences in archaeological data. *Journal of Archaeological Method and Theory*, 24, 640-658.
- Schmidt, P., Blessing, M., Rageot, M., Iovita, R., Pfleging, J., Nickel, K. G., ... & Tennie, C. (2019). Birch tar production does not prove Neanderthal behavioral complexity. *Proceedings of the National Academy of Sciences*, 116(36), 17707-17711.
- Schmidt, P., Buck, G., Berthold, C., Lauer, C., & Nickel, K. G. (2019). The mechanical properties of heat-treated rocks: a comparison between chert and silcrete. *Archaeological and Anthropological Sciences*, 11(6), 2489-2506.
- Schoville, B. J., Wilkins, J., Ritzman, T., Oestmo, S., Brown, K. S., 2017. The performance of heat-treated silcrete backed pieces in actualistic and controlled complex projectile experiments. *Journal of Archaeological Science: Reports*, 14, 302-317.
- Sitton, J., Story, B., Buchanan, B., Eren, M. I., (2020). Tip cross-sectional geometry predicts the penetration depth of stone-tipped projectiles. *Scientific Reports*, 10(1), 1-9.
- Souyouzoglou-Haywood, C., O'Sullivan, A. (Eds.) 2019. *Experimental Archaeology: Making, Understanding, Story-Telling*. Archaeopress Archaeology.
- Speer, C. A., 2018. Using Porcelain Replicas for Precision Control in Flintknapping Experiments. *Advances in Archaeological Practice*, 6, 72-81.
- Stout, D. 2005. The social and cultural context of stone-knapping skill acquisition. In: Roux, V., Bril, B. (Eds.), *Stone Knapping: The Necessary Conditions for a Uniquely Hominin Behaviour*, McDonald Institute for Archaeological Research. Cambridge, pp. 331-340.
- Stout, D., Hecht, E., Khreisheh, N., Bradley, B., Chaminade, T., 2015. Cognitive demands of Lower Paleolithic toolmaking. *PLoS ONE*, 10, e0121804.
- Surovell, T. A., Toohey, J. L., Myers, A. D., LaBelle, J. M., Ahern, J. C., Reisig, B. 2017. The end of archaeological discovery. *American Antiquity*, 82, 288-300.
- Thomas, D. H., 1986. Points on points: a reply to Flenniken and Raymond. *American Antiquity*, 51, 619-627.
- Tsirk, A. 2014. *Fractures in Knapping*. Archaeopress, Oxford.
- Whittaker, J. C., 1994. *Flintknapping: Making and Understanding Stone Tools*. University of Texas Press, Austin.

Whittaker, J. C., 2004. *American Flintknappers: Stone Age Art in the Age of Computers*. University of Texas Press, Austin.

Whittaker, J. C. 2010. Weapon trials: the atlatl and experiments in hunting technology. In: Ferguson, J. (Ed.), *Designing experimental research in archaeology: Examining Technology Through Production and Use*, University Press of Colorado, Boulder, pp. 195-224.

Wilkins, J., Schoville, B. J., Brown, K. S., 2014. An experimental investigation of the functional hypothesis and evolutionary advantage of stone-tipped spears. *PLoS ONE*, 9(8), e104514.

Wilson, M., Perrone, A., Smith, H., Norris, D., Pargeter, J., Eren, M. I. 2021. Modern thermoplastic (hot glue) versus organic-based adhesives and haft bond failure rate in experimental prehistoric ballistics. *International Journal of Adhesion and Adhesives*, 104, 102717.

Figure Captions

Figure 1. Clovis (a), Madison (b), and Side-Notched (c) stone projectile points knapped by M.I.E.

Figure 2. Reconditioning the molding sand with additions of clay (a), and oil (b), removing reconditioned sand ready for use (c).

Figure 3. Tools used in the loose pattern casting process: flask set cope over drag (a); backing boards (b); sprue cutter (c); strike off bar (d); hand ram (e); parting compound (f); pouring basin mold (g); pouring basin demolding frame (h); riddle (i).

Figure 4. Positioning the patterns in the flask (a), application of parting compound (b), sand being loaded into the riddle (c), hand compacting the facing sand layer (d).

Figure 5. Hand Ram (a), Strike-off Bar (b), Sand wrapped around pattern (c), carving the offset parting line (d).

Figure 6. Completed offset parting line (a), Application of parting compound (b), Cutting the sprue with a sprue cutter (c), Gating system completed (d).

Figure 7. pouring Basin (a), Molds on pouring floor (b), Melting the initial metal charge (c), adding additional metal to the existing melt (d).

Figure 8. Taking the temperature of the molten metal (a), Transferring metal to a pouring ladle (b), pouring molds (c), after molds have had time to cool, moving them to molding bench for shake out (d).

Figure 9. Castings at shake out (a), Cooling castings in water (b), cold castings (c), cleaning up the excess metal from the gating system (d).

Figure 10. The hafted aluminum and stone projectile points.

Figure 11. The Clovis (a), Madison (b), and Side-Notched (c) aluminum and stone points. We also turned a plastic cast (d) “back into stone”. The aluminum points in this image have not yet been cleaned up or dremmeled to remove any extraneous aluminum.

Figure 12. Posterior distributions of aluminum versus stone penetration depths (cm) for Clovis, Side-Notched, and Madison projectile points. Aluminum results on top row and stone on bottom row. 95% credible intervals are shaded, and the vertical line is the median of the posterior distribution.

Figure 13. Posterior distributions of aluminum versus stone penetration depths (cm) for Side-Notched projectile points for the second and third trials. Aluminum results on top row and stone on bottom row. 95% credible intervals are shaded, and the vertical line is the median of the posterior distribution.

Figure 14. Arrows are indicating spots of lashing “roughness” present on the aluminum Side-Notched point, but absent on its stone counterpart.

ACCEPTED MANUSCRIPT