

Copying error, evolution, and phylogenetic signal in artifactual traditions: an experimental approach using “model artifacts”

Kerstin Schillinger¹, Alex Mesoudi², and Stephen J. Lycett^{1*}

¹Department of Anthropology, Laboratory of Evolutionary Anthropology, University at Buffalo, SUNY, 380 MFAC-Ellicott Complex, Amherst, NY 14261, USA

²Human Biological and Cultural Evolution Group, Department of Biosciences, College of Life and Environmental Sciences, University of Exeter, Penryn Campus, Cornwall, TR10 9FE, UK

*Corresponding author: sjlycett@buffalo.edu

Abstract

Spatio-temporal patterns of artifactual variation are increasingly being studied via the explicit application of cultural evolutionary theory and methods. Such broad-scale (macroevolutionary) patterns are mediated, however, by a series of small-scale (microevolutionary) processes that occur at the level of individual artifacts, and individual artifact users and producers. Within experimental biology, “model organisms” have played a crucial role in understanding the role of fundamental microevolutionary processes, such as mutation and the inheritance of variation, in respect to macroevolutionary patterns. There has, however, been little equivalent laboratory work to better understand how microevolutionary processes influence macroevolutionary patterns in artifacts and their analysis. Here, we adopt a “model artifact” approach to experimentally study the issues of copy error (mutation) and resultant phylogenetic signal in artifact traditions. We used morphometric procedures to examine shape copying error rates in our “model artifacts.” We first established experimentally that statistically different rates of copying error (mutation) could be induced when participants used two different types of shaping tool to produce copies of foam “artifacts.” Using this as a baseline, we then tested whether these differing mutation rates led to differing phylogenetic signal and accuracy in two separate experimental transmission chains (lineages), involving participants copying the previous participant’s artifact. The analysis demonstrated that phylogenetic reconstruction is more accurate in artifactual lineages where copying error is demonstrably lower. Such results demonstrate how fidelity of transmission impacts directly on the evolution of technological traditions and their empirical analysis. In particular, these results highlight that differing contexts of cultural transmission relating to fidelity might lead to differing patterns of resolution within reconstructed evolutionary sequences. Overall, these analyses demonstrate the importance of a “model artifact” approach in discussions of cultural evolution, equivalent in importance to the use of model organisms in evolutionary biology in order to better understand

fundamental microevolutionary processes of direct relevance to macroevolutionary archaeological patterns.

Introduction

The importance of understanding artifactual traits as part of dynamic, historical systems that are mediated via social learning pathways within communities, has become increasingly apparent (e.g., Eerkens and Lipo, 2007; Shennan, 2011; Jordan, 2015; Lycett, 2015a). Changing patterns of artifactual variation, whether measured in discrete or continuous terms, have the capacity to reflect—at least in part—human decisions, biases, and other relevant historical factors (e.g., drift) over varying scales of space and time (Clarke, 1968; O’Brien and Lyman, 2000; Shennan, 2000; Eerkens and Lipo, 2005; Buchanan and Hamilton, 2009; Mesoudi, 2011; Okumura and Araujo, 2014; Eren et al. 2015; Lycett, 2015b). In other words, patterns in the archaeological record have the capacity to provide insights into the processes underlying *evolving* information systems that comprise human cultures.

Within such a system, the potential for continuity is provided by social learning, which can be defined simply as the learning of something new by watching another individual’s behavior or observing the outcomes of that behavior, such as artifacts (Heyes, 1994). Such a simple definition tends to mask, however, that “social learning” related to artifactual production and use, may potentially take place via a variety of quite distinct learning mechanisms, such as emulation, imitation, or teaching (Lycett, 2015b). However, regardless of how it takes place, social learning has the capacity to create “lineages” or “traditions” in artifactual data, visible over time and space (Clarke, 1968; O’Brien and Lyman, 2000). Such factors have led to an increased use of evolutionary theory being applied to archaeological data over the last decade and, equally importantly, the application of specific analytical techniques drawn from evolutionary biology, such as phylogenetic analysis and formal statistical models from population genetics (for reviews see e.g., Eerkens and Lipo, 2007; Mesoudi, 2011; Shennan, 2011; Lycett 2015b).

As part of these considerations, the role of copying errors that occur in artifactual traditions and their implications for subsequent elements of the evolutionary process have received increased attention (Eerkens, 2000; Eerkens and Lipo, 2005; Hamilton and Buchanan, 2009; Kempe et al., 2012; Schillinger et al., 2014a, 2014b, 2015; Lycett et al., 2016). Copying errors resulting from imperfect replication of cultural entities are an important means by which new variation might be generated, which then might themselves be copied by others (Eerkens and Lipo, 2007; Lycett 2015b). As Eerkens and Lipo (2005: 317) put it, copying errors can be a potent source of cultural

variation which then acts as the “raw material” for subsequent evolution (i.e., descent with modification). Hence, just as is the case with the process of mutation in genetic evolution, this source of variation is a potentially important factor in culturally mediated artifact systems over the longer term. More specifically, copying errors that occur at the level of the individual artifact manufacturer are a *microevolutionary* consideration, which might ultimately come to have an influence on wider patterns that are more strikingly visible at the *macroevolutionary* scale in artifactual traditions. In recent years it has been demonstrated that memory effects (Eerkens, 2000), perceptual errors (Kempe et al., 2012), manufacturing processes (Schillinger et al., 2014a), time constraints (Schillinger et al., 2014b), and mode of social learning (Schillinger et al., 2015) might all influence the rate of mutation in artifact traditions. Despite an increased recognition that the study of microevolutionary factors is directly relevant to the evolutionary analysis of artifactual traditions, much remains to be done in order to understand precisely how these small-scale factors relate to evolutionary processes and larger-scale archaeological patterns. For instance, what effect does a relatively higher or lower copy error (i.e., “mutation”) rate have on the evolution of artifactual lineages and their analysis? Indeed, although considerable attention has been given to the issue of potentially phylogeny-obscuring effects such as between-group rather than within-group transmission in artifactual traditions (e.g., Jordan and Shennan, 2003; Tehrani 2011; Tehrani et al., 2010), far less attention has been given to the role of learning errors in craft traditions in terms of their effects on phylogenetic patterns.

Within evolutionary biology, profound progress over several decades has been made in the understanding of fundamental microevolutionary processes, such as mutation, via laboratory experiments involving so-called “model organisms” (Elena and Lenski, 2003; Futuyma and Bennett, 2009; Garland and Rose, 2009; Mueller, 2009; Bataillon et al., 2013). Model organisms are valuable in this respect because they enable a more secure understanding of phenomena of wide interest, from seemingly discrete, but highly controlled, laboratory experiments. Commonly used model organisms, such as fruit flies (*Drosophila* spp.) or bacteria (e.g., *Escherichia coli*), tend to have a variety of characteristics that make them particularly amenable for use in such experiments, including economy, relative simplicity, speed of reproduction, and manipulability (e.g., Elena and Lenski, 2003; Greenspan, 2004; Ashburner et al., 2005). The most suitable model organisms thus display some of the complexities of the phenomenon of general interest, yet are not so complex that they are unwieldy in experimental settings and thus enable the

precise study of discrete factors and processes relevant in the context of wider evolutionary questions. Given that artifacts are the product of cultural evolutionary processes, we have elsewhere argued that simple laboratory experiments that replicate certain aspects of artifactual form (e.g., their size and/or shape) make a particularly useful subject of study for similar reasons (Schillinger et al., 2014a; Lycett et al., 2015). In other words, similarly to the use of such experiments in evolutionary biology, the examination of “model artifacts” in controlled laboratory settings can help shed light on fundamental microevolutionary processes that are directly relevant to issues that must be considered when examining wide-scale and long-term patterns in the archaeological record (Eerkens, 2000; Mesoudi and O’Brien, 2008; Kempe et al., 2012; Schillinger et al., 2014a, 2014b, 2015, 2016; Lycett et al., 2016).

Here, we focus on the role of copy error (i.e., unintentional artifactual mutation) on the evolutionary process using a model artifact experimental approach. Our study specifically focused on the implications that varying rates of copying error (i.e., differing rates of copying fidelity) might have on subsequent aspects of the evolutionary process in artifact systems. Either within biological evolutionary systems or artifactual systems, transmission fidelity will inevitably create lineages of related entities (either comprised of individual animals, for example, or artifacts) that thus bear certain amounts of external resemblance to each other due to this shared ancestry (Darwin, 1859; Clarke, 1968; O’Brien and Lyman, 2000; Mesoudi, 2011). Equally, however, the introduction of new variation into such inheritance systems via imperfect replication/transmission mechanisms, for example, will lead to a divergence away from existing patterns of variation and thus divergence of similarity (Darwin, 1859; O’Brien and Lyman, 2000; Mesoudi et al., 2004). Our study focused on what varying rates of copying error in artifactual lineages might imply for such patterns of descent and their empirical analysis. This is not to deny the importance of intentional, directed change in cultural evolution (and the presence of such intentional change does not invalidate cultural evolutionary approaches [Mesoudi, 2008]). However, even where intentional change *is* present, it is necessary to understand the effect of purely unintentional copy-error-based change.

The extent to which similarity and historical relatedness may correlate with each other can obviously be affected by underlying microevolutionary processes including the extent of fidelity (replication accuracy) prevalent within a given situation (Hansen and Martins, 1996). Hence, the

degree of correlation between genealogical relationships of artifacts and their measurable properties may influence the extent to which evolutionarily focused analyses (and methods) recover meaningful historical information from the archaeological record. Here, we experimentally examine the extent to which differing mutation rates in artifactual lineages might influence the extent of correspondence between historical relatedness (i.e., genealogical relationship) in artifactual traditions and their measurable properties. As we have noted, in recent years a range of techniques and methods drawn from evolutionary biology, including phylogenetic methods and population genetic models, have been applied to artifactual datasets to examine evolutionary questions. Phylogenetic analyses are one means of examining potential genealogical relationships between artifacts and assemblages based on their observable (i.e., physical) properties and attributes (e.g., O'Brien et al., 2001; Jordan and Shennan, 2003; Harmon et al., 2006; Lycett, 2007; Buchanan and Collard, 2008; Cochrane, 2008, 2013; Riede, 2008; Jordan and O'Neill, 2010; Tehrani, 2011; Jennings and Waters, 2014; Jordan, 2015; Prentiss et al., 2015; Mendoza-Straffon, 2016).

Our study used phylogenetic (maximum parsimony) methods to examine how statistically different rates of copying error in two model artifact “traditions” might influence the extent to which their (known) genealogical relationships could accurately be recovered. Interestingly, within evolutionary biology, one use of model organisms has been to determine the extent to which phylogenetic methods might recover genealogical relationships in laboratory bred (pedigree) lineages (Atchley and Fitch, 1991; Hillis et al., 1992). Spencer et al. (2004) also used copies of written text produced within an experimental context to examine the applicability of phylogenetic methods to historical documents. To our knowledge, however, the artifactual equivalent of such studies, whereby “artifact pedigrees” of known relationship are generated under controlled laboratory conditions and then used to determine the extent to which analytical methods might accurately recover those relationships, have not been undertaken. Hence, one aim of our study was to undertake such an assessment. However, as we have noted, a further aim of our analysis was to determine the extent to which contrasting rates of copying error might influence the relative accuracy of phylogenetic recovery in each case. This broader aim emphasizes that our analyses are focused on understanding the role of specific microevolutionary factors (i.e., copying errors) in wider evolutionary processes relating to artifactual systems, than they are strictly “phylogenetic” (or “tree building”) issues per se.

To accomplish these aims, the experimental study was comprised of two distinct phases. Firstly we asked a series of volunteers to copy the shape of a target-model foam “handaxe” from a foam block using either a plastic knife or a metallic peeler as a shaping tool. As is the case with many shaped artifacts (pots, baskets, projectile points, etc.), the effective production of stone handaxes requires levels of skill and experience that are built over months, if not years, of practice (Edwards, 2001). Conversely, foam “handaxes” are easily manufactured, requiring no specialist skills or knowledge. However, while “handaxe” manufacture from foam blocks is a simpler task compared to real stone-knapping, it still requires the manufacturer to accurately copy integrated and multivariate 3D shape properties, such as relative length, breadth, and thickness variables, and overall outline shape, in order to accurately replicate overall artifactual form. Thus, as with model organisms these “model artifacts” are experimentally useful in respect to gaining insights into factors relevant to a far wider range of artifactual situations. Indeed, artifactual shape, invariably has socio-functional and techno-functional relevance (*sensu* Schiffer and Skibo, 1987) in cultural systems. The aim of this initial experimental phase was to establish that use of these two shaping tools (i.e., the plastic knife versus the metallic peeler) resulted in statistically different rates of copying error in the artifact copies produced.

Having established this statistical relationship between manufacturing tool and rate of copying error, we then undertook our main experimental component. We implemented two separate experimental conditions using a transmission chain method (Mesoudi, 2007). Often attributed to Bartlett (1932), who pioneered a series of experiments in social psychology using transmission chains (Mesoudi, 2011), versions of this experimental procedure were, however, used somewhat earlier by those interested in cultural transmission and indeed artifactual change (e.g., Balfour, 1893). The transmission chain method involves passing a design, words, or idea along a chain of participants from one to the other in succession, thus creating an experimental “lineage” that may change according to the introduction and transmission of new variation with each succeeding transmission event. In our experiment, we implemented two such transmission chains, whereupon participants in each chain used the plastic knife to produce foam handaxes in one case, and the metallic peeler in the alternative condition. Hence, the first participants in each chain were provided with the initial target foam handaxe, and then each participant was given the previous participant’s artifact to copy in each respective case. This method thus provided two artifact lineages, whereupon the genealogical relationships of the artifacts in each chain were

known precisely. We then used morphometric procedures to obtain shape data for the artifacts in each chain. Given that use of the plastic-knife tool is known to result in statistically greater copying error compared to use of the metallic-peeler tool, we predicted that phylogenetic signal (i.e., correspondence between artifact morphology and accuracy of genealogical information) would be greater in the chain that used the metallic-peeler tool compared to the plastic-knife tool. We then tested this prediction using formal phylogenetic (maximum parsimony) methods. We were also able to examine the role that differing rates of copying error had on the overall evolutionary pattern within the two separate artifactual lineages, thus providing new insights into the overall evolutionary role of copying error in the case of archaeologically observable traditions.

Materials and Methods

Establishing baseline copying error using the two different manufacturing tools

Prior to the main (phylogenetic) analyses we first established that use of the plastic knife resulted in statistically greater amounts of copying error compared to using the metallic peeler. To achieve this, we recruited a total of 60 volunteers. We asked basic demographic information of the participants including their sex and age group. These participants were undergraduate and postgraduate students, recruited via advertisement. Informed consent was obtained from all volunteers. 30 participants were female (aged between 18 to a maximum of 44 years based on self-reported categories) and 30 were male (aged between 18 to a maximum of 34 years based on self-reported categories). A financial compensation of £4 (~\$6 US) was paid to each participant as a reward for their involvement in the experiment.

The experiment required all participants to copy the same three-dimensional foam “target form,” modeled after the shape of an Acheulean handaxe (Figure 1). The target model and all subsequent copies of it produced by the participants were manufactured from standardized blocks of foam. These high-density foam blocks (OASIS DRY SEC foam) are of a type used by professional florists to securely hold artificial flowers. The material is robust to handling, but designed to be malleable so that it can be easily cut and modified into desired sizes and shapes using commonplace implements such as knives and scissors. The machine-cut foam blocks were acquired from the manufacturer with standardized dimensions (22.3cm length × 11cm width ×

7.8cm thickness), which allowed for the starting conditions during the experiment to be identically replicated across all of the participants.

Replica “handaxes” were manufactured from the foam blocks using one of two different tools. One foam manipulation tool was a simple plastic table-knife (Figure 2a). The second tool was a metallic vegetable-peeler (REX Swiss Quality peeler). As shown in Figure 2b, the metallic peeler has a movable blade positioned perpendicular to the handle. Both the plastic knife and the metallic peeler were specifically chosen for the experiment given their equal suitability for both left and right handed use, thus controlling for this handedness; however, as an additional control we also ensured that the proportion of right and left handed participants in each condition was statistically identical ($\chi^2 = 0.5769$, $p = 0.447$). Participants were additionally provided with a lab coat, dust mask, and safety goggles to guard from foam dust.

The participants in this experiment were divided equally between the two experimental conditions (i.e., the “metallic-peeler condition” and the “plastic-knife condition”) such that there were 30 participants in each group. Studies in psychology have shown that human males and females may differ (statistically) in their performance during tests of spatial perception and mental rotation tasks involving 2D objects (Halpern, 2000; Linn and Peterson, 1986; Voyer et al., 1995; Wynn et al., 1996) although it should be noted that such sex-related effects are also known to reduce substantially whenever 3D objects (rather than 2D representations) are used in such cognitive tests (Robert and Chevrier, 2003). For the purposes of precaution, however, we specifically controlled for sex differences by dividing participants between the alternate tool conditions such that there were equal numbers (i.e., 15 females and 15 males) in each condition.

While participants in both conditions were asked to faithfully copy the overall shape and form of the target foam handaxe, it was clearly specified that the priority goal was to copy the *shape* of the target form. To enhance motivation of participants during the experiment, each participant was informed that whoever most accurately copied the target (model) handaxe would receive a book voucher to the value of £20, in addition to the £4 compensation. All participants undertook the experiment individually and no participant saw another person make their particular handaxe. Prior to commencing the copying task, each participant read an instruction sheet and then was handed the “target” foam handaxe. Each participant was then instructed to inspect this target

form from all sides and angles for one minute. It should be noted that the model remained with the participants throughout the task and they were permitted to make direct comparisons with their own copy at any time, which controlled for potential memory effects (e.g., Eerikens, 2000). Participants were then placed at a table and provided with a standardized foam block and one of the two alternative types of manipulation tool. Participants were given a 20 minute timeframe to complete the task. Previous experiments have demonstrated that 20 minutes is sufficient time to effectively complete the replication task (Schillinger et al., 2014b, 2015). Participants were provided with a countdown timer and were verbally reminded of the remaining time at five minute intervals. Participants were also permitted to wear spectacles or contact lenses, where necessary, for close-up tasks to avoid biases resulting from visual disparities between participants.

A total of 42 morphometric variables were recorded on each handaxe replica produced by each participant. Of these, 28 measurements were taken from the plan-view perspective (Supplemental Figure 1a) and 14 measurements were obtained from the profile-view perspective (Supplemental Figure 1b). The measurements were digitally recorded from photographic images that were imported into the freely available morphometric software *tpsDig* v2.16 (Rohlf, 2010). Standardized photographic images of the plan- and profile-view perspective were taken using a Fujifilm DSLR camera (30x zoom lens: 24-720 mm) fixed to a copystand. Each handaxe replica was positioned according to a standardized orientation protocol to obtain maximally homologous measurements (Supplemental Material). The measurements for each handaxe replica were recorded by superimposing a digital grid on each photographic image that defined the 42 bilateral and lateral measurements. The procedure for obtaining morphometric shape data has been fully described in previous experimental settings (see e.g., Schillinger et al., 2014a, 2014b), and further details regarding the orientation protocol are provided in the electronic supplementary information.

Since the focus of the analysis was on *shape* copy-error rates, data from the 42 morphometric variables were size-adjusted via use of the geometric mean method prior to analysis (Jungers et al., 1995; Lycett et al., 2006). This method effectively removes size (scaling) variation between objects while retaining relevant shape data (Falsetti et al., 1993; Jungers et al., 1995). The geometric mean derived from a series of n variables (e.g., a_1 to a_n) can be calculated using the

formula: $\sqrt[n]{a_1 \times a_2 \times a_3 \times \dots \times a_n}$ (Sokal and Rohlf, 1995: 43). Size-adjustment of the raw data proceeds on a specimen-by-specimen basis, and involves simply dividing each variable of an individual specimen, in turn, by the geometric mean of all the variables for that particular specimen; this procedure is then repeated for all specimens.

To determine statistically whether copying error was significantly different between the plastic-knife group and the metallic-peeler group, we first calculated copying error for each individual's foam handaxe relative to the target model. To accomplish this, copying error rates were extracted from the size-adjusted data set using a straightforward two-step method. Firstly, the 42 size-adjusted variables for each handaxe replica were subtracted from the equivalent 42 variables of the model target form. Secondly, mean shape error rates were computed for each of the 42 morphometric variables across the 30 handaxe copies produced in each of the two tool groups. These 42 computed error-rates for each experimental condition were subsequently compared via statistical analysis.

In order to determine whether the two different manufacturing tools generated statistically distinct levels of shape copying error, a Mann-Whitney U test was undertaken (Monte Carlo, 10,000 random assignments, $\alpha = 0.05$) using PAST v3.08 (Hammer et al., 2001). This procedure demonstrated that participants in the metallic-peeler condition generated an average shape copying error rate of 0.121 in the size-adjusted data (SD = 0.067). In the plastic knife condition, average participant copy-error rates were 0.137 (SD = 0.047). Statistical comparison of these error rates demonstrated that they were significantly different ($U = 629.5$, Monte Carlo $p = 0.023$). Hence, this analysis demonstrated that using a plastic knife to copy a target foam handaxe, on average, generated a significantly higher rate of shape copying-error compared to using a metallic peeler.

Transmission chain conditions: the “plastic-knife chain” versus the “metallic-peeler chain”

Having established that use of plastic knives results in a significantly greater rate of copying error compared to using metallic peelers, we then instigated an experiment that implemented a “transmission chain” method. This effectively created two sets of artifactual “pedigrees” (or

artifact “lineages”), whereupon genealogical relationships between all handaxes were known based on their position in their particular chain.

A total of 30 new male participants (age range 18 to a maximum of 34 years) took part in the experiment, divided equally between the two manufacturing (i.e., tool group) conditions. Again, the majority of participants were undergraduate and postgraduate students, recruited via advertisement. All instructions and experimental conditions provided to participants were the same as those used for the baseline (copying error) experiment described previously. The first participant in each condition received the same initial “target” model as in the previous baseline experiment. Thereafter, each participant in the chain received the previous participant’s handaxe replica as their target model to copy. Participants did not observe the previous participant manufacturing their model and were only provided with the artifact “target” as the basis for copying, consistent with the conditions in the baseline copying-error experiment. As before, to enhance motivation levels in the participants, they were informed that the participant who most accurately copied the foam handaxe would receive a book voucher to the value of £20, in addition to the £4 compensation.

Participants in the “plastic-knife chain” used a plastic knife to manufacture their copies, while participants in the “metallic-peeler chain” were issued with a metallic peeler as their manufacturing tool. This experimental procedure resulted in two sets of transmission-chain handaxe replicas for each manufacturing condition (Figures 3 and 4). A total of 16 handaxes comprised each chain, one for each of the participants in each condition plus the initial “target” model provided to the first participant in each chain (Figures 3 and 4). Thereafter, morphometric data in plan- and profile-views (total 42 variables) were obtained for all handaxe replicas using the procedure described above. Again, data were size-adjusted via the geometric mean method prior to analysis in order to maximize shape information.

Phylogenetic reconstruction via parsimony

Cladistic analysis is based on finding the phylogenetic tree (i.e., genealogy) that most parsimoniously explains the evolution of the inputted data, which is why it is also often referred to as “parsimony analysis” (Swofford and Sullivan, 2009). Computationally, this method finds the tree (or trees) that require the minimum number of evolutionary steps to account for the data.

Cladistic analysis achieves this by searching for trees that contain shared patterns of character change between taxa (i.e., groupings that contain “shared-derived” patterns of character evolution) while simultaneously minimizing the number of evolutionary convergences (Kitching et al., 1998; Schuh and Brower, 2009). The direction of evolutionary change (or “character polarity”) may be estimated via reference to an “outgroup,” which is simply a point of reference in the dataset that is known to have branched away from the main group at an earlier point in the evolutionary sequence. Several accessible introductions to cladistic methodology are now available that explain specific aspects of the method in greater detail (e.g., Kitching et al., 1998; Lipscomb, 1998; McLennan and Brooks, 2001), including some written specifically with archaeologists in mind (e.g., O’Brien et al., 2001; O’Brien and Lyman, 2003). Such methods have been used in a range of artifact-focused studies within the last few years (e.g., Harmon et al., 2006; Lycett, 2007; Buchanan and Collard, 2008; Cochrane, 2008, 2013; Jordan and O’Neill, 2010; Tehrani et al., 2010; Tehrani, 2011; Jennings and Waters, 2014; O’Brien et al., 2014; Jordan, 2015; Prentiss et al., 2015).

Parsimony analysis was conducted using the freely available software TNT (Goloboff et al., 2008). An advantage of the TNT software is that it facilitates direct analysis of morphometric (i.e., continuous) data without having to resort to converting the characters into discrete states (Goloboff et al., 2006, 2008). The tree bisection–reconnection (TBR) algorithm was used in searching for the shortest, most-parsimonious trees. Although the TBR algorithm is the most costly in terms of computation time, it is also the most effective in recovering the most parsimonious cladograms (Kitching et al., 1998:46).

Given that here the exact sequence of evolution is known, we can directly compare this known sequence to that reconstructed via the parsimony analysis for each data set in order to readily assess overall accuracy of phylogenetic reconstruction in each case.

Retention Index (RI) comparison

A further method of examining the phylogenetic signal in our two datasets is to inspect a commonly used descriptive statistic for cladistic trees referred to as the “ensemble Retention Index” or RI value (Kitching et al., 1998: 97). When the data perfectly fit a particular phylogenetic tree (i.e., no convergent instances of character evolution are required) the characters are considered to be fully consistent with that particular tree shape. Conversely, “homoplastic” or

convergent character changes (i.e., the character state changes more than once at different points in the tree) are the reason why data sets are typically less than fully consistent with a given tree (Goloboff, 1991; Kitching et al., 1998). The ensemble retention index (RI) is a tree statistic commonly used to describe the amount of homoplasy (i.e., convergence) in a given tree (Farris, 1989a, 1989b). The RI expresses homoplasy relative to the amount of character change that is retained during the hierarchical process of branching as revealed by the most parsimonious cladogram (Kitching et al., 1998; Lipscomb, 1998). In other words, the RI for a given tree (cladogram) is a measure of how much evolutionary convergence there is within that tree; the lower the number of convergences, the stronger the data support the overall tree structure. RI values range between 0–1, with values close to 1 indicating a strong fit between the dataset and the shape of the tree, and values close to 0 indicating a poorer fit between the data and the shape of the tree (i.e., a higher number of convergences). Indeed, Nunn et al. (2010) have argued on the basis of computer simulation that higher mutation rates can negatively impact RI scores, indicating the suitability of using this measure given our goals here. Conveniently, the RI statistic is insensitive to differences between the dimensions of different character matrices, thus allowing values generated from different data sets to be compared directly. If copying error negatively influences phylogenetic signal, then we can reasonably predict that the RI score for the plastic-knife data will be lower than that for the metallic-peeler data.

Bootstrap analyses

In phylogenetic analysis, “bootstrapping” is a method for assessing the level of support for individual clades within a given cladogram (Felsenstein, 1985; Kitching et al., 1998). The bootstrap procedure randomly samples characters with replacement, and forms a pseudoreplicate matrix with the same dimensions (number of characters and character states) as the original data matrix. This resampling procedure is undertaken a large number of times (e.g., creating 1000 pseudoreplicate matrices). Thereafter, parsimony analysis is applied to each pseudoreplicate matrix, with results typically presented in the form of a majority-rule consensus tree as a means of displaying the proportion of resampled data sets that support individual nodes. In other words, relationships that are relatively robust to randomization will have higher bootstrap support than those that are more sensitive to the randomization procedure, which will have relatively lower values. Typically, the results of bootstrap analysis are displayed in the form of a 50% majority-

rule consensus tree, whereupon groupings represented with frequencies greater than 50% in the bootstrap analysis are indicated, while those below 50% not typically reported (Kitching et al., 1998: 129). Here, to obtain bootstrap values below 50% in addition to those higher than this value, we used the frequency differences (GC) option in TNT with cut-off set at 5% (Goloboff et al., 2008). We undertook 1000 pseudoreplicates of the original data matrix. Thereafter, we undertook a Mann-Whitney U-test of bootstrap values for each internal node, comparing values for the plastic-knife chain against the metallic-peeler chain. If copying error influences phylogenetic signal, we can expect that values for the metallic-peeler chain will be statistically higher than for the plastic-knife chain.

Results

Phylogenetic reconstruction via parsimony

Figure 5a shows the (single) phylogenetic tree produced via parsimony analysis of data from the plastic knife transmission chain. The first two steps of the known phylogenetic sequence are reconstructed correctly (Figure 5a). The analysis does, however, incorrectly position handaxes in later stages of the transmission chain, misplacing PK4, PK5, PK8, and PK9, and incorrectly indicating the last two handaxes (PK15 and PK16) share a closer relationship with earlier stages of the chain than they do in reality (Figure 5a).

Figure 5b shows the equivalent (single) phylogenetic tree for the metallic peeler transmission chain. Some of the earliest steps in the transmission chain are incorrectly positioned, which is particularly noticeable in the case of MP4 and MP5 that are incorrectly placed at an earlier stage in the cladogram than they should be based on the known sequence. From the sixth step in the tree, however, phylogenetic relationships are highly consistent with the known sequence of the transmission chain (Figure 5b).

As a basic point of comparison between these two trees, it is notable that the plastic knife tree shows greater overall inaccuracy (six misplaced handaxes) compared to the metallic peeler (four misplaced handaxes). It can also be observed that the metallic peeler tree shows least accuracy in the earliest branches but shows greater accuracy in the later phases of the chain.

Retention Index (RI) comparison

The RI value for the phylogenetic tree produced for the plastic knife transmission chain data was 0.66. Conversely, the RI value for the metallic peeler transmission chain tree was 0.81. Hence, these retention index values indicate that there is higher phylogenetic signal in the metallic peeler data compared with the plastic knife data.

Bootstrap analyses

Bootstrap analysis of the plastic knife data produced a mean bootstrap value of 43.8% (median = 28.5%) (Figure 6a). It is notable that lower bootstrap values tend to occur in the latest stages of the phylogenetic tree in the plastic-knife condition (Figure 6a). Bootstrap analysis of the data for the metallic-peeler transmission chain produced a mean value of 77.1% (median = 88.5%) (Figure 6b).

Figure 7 compares the bootstrap values for the plastic-knife transmission chain and the metallic peeler transmission chains via box plots, which emphasizes the differences in the bootstrap values between these two experimental conditions. The Mann-Whitney U -test indicated that these bootstrap values are statistically different in the two experimental conditions ($U = 34.5$; Monte Carlo $p = 0.029$). Hence, overall these results indicate that phylogenetic signal is higher in the metallic peeler transmission chain compared with the plastic knife transmission chain.

Discussion

Copying errors inevitably arise during the manufacture of hand-crafted artifacts (Clarke, 1968; Eerkens, 2000; Kempe et al., 2012; Schillinger et al., 2014a). The role that differing “mutation” rates in the physical attributes of artifacts might have on their subsequent evolution is, however, an underexplored avenue of research. Such a gap in current endeavors is potentially problematic since a full understanding of small-scale, microevolutionary processes, is essential if we are to fully analyze and understand macroevolutionary patterns attested in spatio-temporal variation of artifacts within the archaeological record. Here, using a “model artifact” approach, we have experimentally examined the role of demonstrably different mutation rates on subsequent evolutionary processes. Moreover, we have examined the effect of these differing mutation rates on the extent of phylogenetic signal that is recoverable solely via their measurable physical properties.

Our overall research design was comprised of two distinct phases. The first phase involved running an initial experiment in order to demonstrate that use of two different manufacturing tools (a plastic knife versus a metallic peeler) resulted in statistically different levels of accuracy in copies of foam artifacts. Having established this, we were then able to use this information to derive predictions regarding phylogenetic signal in two separate “transmission chains” produced in our primary experiment, whereupon participants in each respective chain exclusively used one of the two shaping tools to produce their artifacts. Specifically, we predicted that phylogenetic signal would be lower in cases where the rate of copying error (i.e., artifact mutation) is known to be higher. We then tested this prediction by applying formal phylogenetic (cladistic) methods to the artifactual data derived from our two experimental “model artifact” transmission chains.

Our phylogenetic analyses indicated a general level of correspondence between the morphometric data obtained from the artifacts and their known position in the genealogical sequence. That is, in the case of data from both the plastic-knife condition and the metallic-peeler condition, artifacts produced earlier in the chains were generally placed earlier in the phylogenetic sequence, whereas artifacts produced later in the chain were generally placed at later points in the branching sequence. In other words, in the absence of a known genealogical sequence for these artifacts, phylogenetic analysis of morphometric data relating to their physical attributes would reveal useful evolutionary information. As we have noted, one of the uses of laboratory “model organisms” in evolutionary biology has been the testing of phylogenetic methods on pedigrees of known relationship (e.g., Atchley and Fitch, 1991; Hillis et al., 1992). Our analyses of the archaeological equivalent of laboratory generated lineages of “model artifacts,” thus support the use of phylogenetic methods in obtaining important evolutionary information in the case of archaeological data where such historical information would be totally unknown *a priori* (e.g., Harmon et al., 2006; Buchanan and Collard, 2008; Cochrane, 2008, 2013; Riede, 2008; Jordan and O’Neill, 2010; Tehrani, 2011; Jennings and Waters, 2014; O’Brien et al., 2014; Jordan, 2015; Prentiss et al., 2015).

Differences in the extent of phylogenetic signal in each experimental condition were, however, evident in our analyses. In correspondence with theoretical predictions, the increased copy error rate in the plastic-knife chain resulted in a lower degree of phylogenetic signal compared to equivalent data for the metallic-peeler transmission chain. This difference of phylogenetic signal

was evident not only in the higher number of misplaced artifacts in the phylogenetic tree for the plastic-knife data compared with phylogenetic analysis of the equivalent metallic-peeler data, but also in the Retention Index (RI value) for each tree and in the bootstrapping analyses. In the case of the RI values, as predicted, the plastic-knife condition demonstrated a poorer fit between data and the tree compared with the metallic-peeler condition. In the case of the bootstrap analyses, values were statistically greater in the case of the metallic-peeler data compared with the plastic-knife data. Hence, these analyses consistently demonstrated that a higher copy error (i.e., mutation) rate in the plastic-knife transmission chain, resulted in lower overall phylogenetic signal compared with the metallic-peeler transmission chain. Given that in recent years it has been experimentally shown that a range of different factors might affect the rate of copy error in artifactual traditions (e.g., Eerkens, 2000; Kempe et al., 2012; Schillinger et al., 2014a, 2014b, 2015) these results thus suggest that the operation of such mechanisms will also influence the extent of evolutionary information that can be obtained in varying archaeological situations. This again emphasizes that an increased understanding of microevolutionary processes (e.g., artifactual mutation rates) is an area that should receive greater research focus.

It should, however, be noted the participants in our two transmission chains were only given the opportunity to copy the previous participant's artifact or what some would term the behavioral "end state" (Whiten et al., 2004), and were not given opportunity to observe the previous participant's behavioral actions that they used during the manufacturing process. In other words, in terms of the specific social learning mechanism used during the copying phase, participants in both of our experimental conditions were using an *emulative* learning mechanism as opposed to a *imitative* learning mechanism (Lycett, 2015b; Schillinger et al., 2015). Recently, we have shown that imitative learning generates statistically lower rates of artifactual copying error compared to emulative learning processes (Schillinger et al., 2015). Hence, in the case of any artifactual tradition attested in the archaeological record where imitative (i.e., action based) learning was used in addition to emulative learning, then phylogenetic signal in artifactual traditions would potentially be greater than indicated in our experimental situation. Nevertheless, we have shown that when the learning process is kept constant (i.e., emulation in both conditions), copying error significantly affects lineage reconstruction.

One further factor that emerged through the results of our experiment is the manner in which varying copying error rates might influence the overall evolutionary process. In the case of the metallic-peeler condition, for instance, where rate of artifactual mutation was comparatively lower, phylogenetic analysis of the data showed greatest inaccuracy during the earliest phases of the transmission chain, but following the initial steps of the chain it was able to reconstruct the known sequence accurately. Conversely, in the case of the plastic-knife condition where mutation rates were relatively higher, inaccuracy during phylogenetic reconstruction occurred at multiple points in the tree. In other words, in the case of the metallic-peeler chain, a relative lack of measurable artifactual variation due to low error rate actually reduced the amount of phylogenetic signal in the earliest part of the tree. However, once variation did occur in the chain, some of this new variation was able to be replicated by the next participant such that the transmission chain followed an evolutionary sequence that phylogenetic analysis was able to detect with reasonable accuracy. Conversely, although variation due to error was occurring earlier in the plastic-knife chain, the new variation was relatively less “transmittable” due to the overall extent of error/mutation which was induced by use of the plastic-knife tool by each participant in that particular chain. This further emphasizes the role of transmission fidelity in cumulative and directional evolutionary change in technological traditions, and indeed, the role of mechanisms that might reduce the occurrence of copying errors in material culture traditions (Tomasello, 1999; Lewis and Laland, 2012; Kempe et al., 2014; Schillinger et al., 2015).

A related point here is that despite many noted limitations with archaeological data, archaeologists frequently have access to a temporal sequence and may know something about the specifics of that sequence (within limits) in terms of dates. When such conditions are available, this means that phylogenetic reconstructions might usefully be compared against temporal data for the same dataset to determine to what extent chronological patterning is also reflected phylogenetically (e.g., Riede, 2009; Lycett, in press). Our bootstrap results indicate that when there is a correspondence between chronology and phylogeny, then changing patterns of phylogenetic signal across a tree (as indicated by bootstrap values) may possibly be related to the extent of transmission fidelity in particular archaeological cases. That is, if there is low phylogenetic resolution in the earliest phases of an archaeological sequence followed by higher phylogenetic resolution in later phases, then this might indicate a tradition that is being transmitted with relatively high fidelity, whereupon it takes longer for detectable evolutionary

change to take effect (i.e., equivalent to our metallic-peeler condition). Conversely, if there is a chrono-phylogenetic sequence where resolution starts relatively high but then deteriorates, then this might indicate a tradition with relatively low transmission fidelity, whereupon new variation is being generated but not transmitted with enough fidelity to result in highly resolved patterns in later stages. We caution, however, that additional tests or independent evidence be used to support any such inference since multiple factors might potentially impact bootstrap values in given circumstances. At the very least, our results reiterate the value of exploring how cultural microevolutionary processes, which are frequently unobservable in the archaeological record, can influence observable macroevolutionary patterns, and in so doing help to build a more robust cultural evolutionary body of theory that better links an understanding of cultural microscale processes to archaeological macroscale patterns.

These results have further implications for the archaeological record. Notably, in archaeological situations where cultural practices lead to an extremely high level of copying fidelity, i.e., effectively resulting in artifactual “clones,” then phylogenetic reconstruction over time may be relatively difficult. One extreme example of such a situation may, for instance, be the case of Japanese sword blades dating from *c.* 1185–1876, which show limited variation during this period due to the strict adherence to metallurgic traditions by swordsmiths at this time (Martin, 2000). Hence, counterintuitively, this implies that in some instances it may be profitable to give greater consideration to artifactual traits that exhibit relatively higher rates of spatio-temporal variation as opposed to those that exhibit minimal variation. Conversely, of course, the results from our plastic-knife chain imply there is a specific “threshold” to this principle, whereupon a rate of variation too high will potentially swamp historical information within the data. Identifying where the limits of such “optimal” levels of variation lie within real artifacts will inevitably require further work. However, given that the empirical measurement of artifactual variation in the archaeological record has long been recognized as a readily accessible component of the available data (e.g., Clarke, 1968), the potential for identifying such issues in archaeological data remains encouragingly high (Eerkens and Lipo, 2007). A further factor to consider here, however, is artifactual mutation that is directly linked to behavioral factors (i.e., is potentially “transmittable” via social learning) compared to mutation that is linked to solely to “environmental” factors (e.g., raw material) (Lycett and von Cramon-Taubadel, 2015). Such a distinction is necessary since this will relate directly to the strength of “evolutionary response”

(Lycett et al., 2016) that may be induced in particular archaeological circumstances and, in turn, the extent of phylogenetic divergence.

Conclusions

It is being increasingly recognized that broad-scale patterns in the archaeological record that are measurable over both spatial and temporal scales may reflect the influence of a range of historical, cultural evolutionary processes. However, such macroevolutionary patterns both begin, and are mediated by, rather less visible *microevolutionary* processes. One such microevolutionary process is the rate of copying error that occurs during social learning, which will result in differing rates of artifactual “mutation” (Eerkens, 2000; Eerkens and Lipo, 2005; Hamilton and Buchanan, 2009; Kempe et al., 2012; Schillinger et al., 2014a, 2014b, 2015). Here, we have demonstrated under laboratory controlled conditions that phylogenetic methods applied to the measurable, physical attributes of artifacts—as might be measured in the case of real artifacts—can yield important evolutionary information. Hence, these results are encouraging for efforts directed toward the recovery of evolutionary information from the archaeological record, even though the physical attributes of artifacts only provide a proxy for many of the behavioral factors that will be the direct source and subject of evolutionarily relevant forces. Moreover, we have identified that specific microevolutionary factors, i.e., differing copying error rates in varying archaeological situations, will also influence the extent to which such analyses may provide accurate historical information. These results, in turn, also raise implications for determining how artifactual mutation rates may guide selection of data that yield the most accurate phylogenetic signal in varying archaeological circumstances.

References

- Ashburner, M., Golic, K., Hawley, R.S., 2005. *Drosophila: A Laboratory Handbook*, second ed. Cold Spring Harbor Laboratory Press, New York.
- Atchley, W.R., Fitch, W.M., 1991. Gene trees and the origins of inbred strains of mice. *Science* 254, 554-558.
- Balfour, H., 1893. *The Evolution of Decorative Art*. MacMillan and Co, New York.
- Bartlett, F.C., 1932. *Remembering*. Macmillan, Oxford.

- Bataillon, T., Joyce, P., Sniegowski, P., 2013. As it happens: current directions in experimental evolution. *Biology Letters* 9, 20120945.
- Buchanan, B., Collard, M., 2008. Phenetics, cladistics, and the search for the Alaskan ancestors of the Paleoindians: a reassessment of relationships among the Clovis, Nenana, and Denali archaeological complexes. *Journal of Archaeological Science* 36, 1683-1694.
- Buchanan, B., Hamilton, M., 2009. A formal test of the origin of variation in North American early Paleoindian projectile points. *American Antiquity* 74, 279-298.
- Clarke, D.L., 1968. *Analytical Archaeology*. Methuen, London.
- Cochrane, E., 2008. Migration and cultural transmission: investigating human movement as explanation for Fijian ceramic change, in: O'Brien, M.J. (Ed.), *Cultural Transmission and Archaeology: Issues and Case Studies*. Society for American Archaeology Press, Washington, D.C., pp. 132-145.
- Cochrane, E.E., Rieth, T.M., Dickinson, W.R., 2013. Plainware ceramics from Sāmoa: insights into ceramic chronology, cultural transmission, and selection among colonizing populations. *Journal of Anthropological Archaeology* 32, 499-510.
- Darwin, C., 1859. *On the Origin of Species*. John Murray, London.
- Edwards, S.W., 2001. A modern knapper's assessment of the technical skills of the Late Acheulean biface workers at Kalambo Falls, in: Clark, J.D. (Ed.), *Kalambo Falls Prehistoric Site: Volume III*. Cambridge University Press, Cambridge, pp. 605-611.
- Eerkens, J.W., 2000. Practice makes within 5% of perfect: visual perception, motor skills, and memory in artifact variation. *Current Anthropology* 41, 663-668.
- Eerkens, J.W., Lipo, C.P., 2005. Cultural transmission, copying errors, and the generation of variation in material culture and the archaeological record. *Journal of Anthropological Archaeology* 24, 316-334.
- Eerkens, J.W., Lipo, C.P., 2007. Cultural transmission theory and the archaeological record: context to understanding variation and temporal changes in material culture. *Journal of Archaeological Research* 15, 239-274.
- Elena, S.F., Lenski, R.E., 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics* 4, 457-469.
- Eren, M.I., Buchanan, B., O'Brien, M.J., 2015. Social learning and technological evolution during the Clovis colonization of the New World. *Journal of Human Evolution* 80, 159-170.
- Falsetti, A.B., Jungers, W.L., Cole III, T.M., 1993. Morphometrics of the Callitrichid forelimb: a case study in size and shape. *International Journal of Primatology* 14, 551-572.
- Farris, J.S., 1989a. The retention index and homoplasy excess. *Systemic Zoology* 38, 406-407.
- Farris, J.S., 1989b. The retention index and the rescaled consistency index. *Cladistics* 5, 417-419.

- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783-791.
- Futuyma, D.J., Bennett, A.F., 2009. The importance of experimental studies in evolutionary biology, in: Garland, Y., Rose, M.R. (Eds.), *Experimental Evolution: Concepts, Methods, and Applications*. University of California Press, Berkeley, pp. 15-30.
- Garland, T., Rose, M.R. (Eds.), 2009. *Experimental Evolution: Concepts, Methods, and Applications*. University of California Press, Berkeley.
- Goloboff, P.A., 1991. Homoplasy and the choice among cladograms. *Cladistics* 7, 215–232.
- Goloboff, P.A., Mattoni, C., Quinteros, Y.S., 2006. Continuous characters analyzed as such. *Cladistics* 22, 589-601.
- Goloboff, P.A., Farris, J.S., Nixon, K.C., 2008. TNT, a free program for phylogenetic analysis. *Cladistics* 24, 774-786.
- Greenspan, R.J., 2004. *Fly Pushing: The Theory and Practice of Drosophila Genetics*. Cold Spring Harbor Laboratory Press, New York.
- Halpern, D.F., 2000. *Sex Differences in Cognitive Abilities*. Lawrence Erlbaum Associates, Mahwah, New Jersey.
- Hamilton, M.J., Buchanan, B., 2009. The accumulation of stochastic copying errors causes drift in culturally transmitted technologies: quantifying Clovis evolutionary dynamics. *Journal of Anthropological Archaeology* 28, 55-69.
- Hammer, Ø., Harper, D.A.T., Ryan, P.D., 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologica-Electronica* 4, 1-9.
- Hansen, T.F., Martins, E.P., 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution*, 50: 1404-1417.
- Harmon, M.J., Van Pool, T.L., Leonard, R.D., Van Pool, C.S., Salter, L.A., 2006. Reconstructing the flow of information across time and space: a phylogenetic analysis of ceramic traditions from prehispanic western and northern Mexico and the American southwest, in: Lipo, C.P., O'Brien, M.J., Collard, M., Shennan, S. (Eds.), *Mapping Our Ancestors: Phylogenetic Approaches in Anthropology and Prehistory*. Aldine Transaction, New Brunswick, NJ, pp. 209-229.
- Heyes, C.M., 1994. Social learning in animals: categories and mechanisms. *Biological Reviews* 69, 207-231.
- Hillis, D.M., Bull, J.L., White, M.E., Badgett, M.R., Molineux, I.J., 1992. Experimental phylogenetics: generation of a known phylogeny. *Science* 255, 589-592.
- Jennings, T.A., Waters, M.R., 2014. Pre-Clovis lithic technology at the Debra L. Friedkin site, Texas: comparisons to Clovis through site-level behavior, technological trait-list, and cladistic analyses. *American Antiquity* 79, 25-44.

- Jordan, P., 2015. *Technology as a Human Social Tradition: Cultural Transmission among Hunter-Gatherers*. University of California Press, Berkeley.
- Jordan, P., Shennan, S., 2003. Cultural transmission, language, and basketry traditions amongst the California Indians. *Journal of Anthropological Archaeology* 22, 42-74.
- Jordan, P., O'Neill, S., 2010. Untangling cultural inheritance: language diversity and long-house architecture on the Pacific Northwest Coast. *Philosophical Transactions of the Royal Society B* 365, 3875-3888.
- Jungers, W.L., Falsetti, A.B., Wall, C.E., 1995. Shape, relative size, and size adjustments in morphometrics. *Yearbook of Physical Anthropology* 38, 137-161.
- Kempe, M., Lycett, S.J., Mesoudi, A., 2012. An experimental test of the accumulated copying error model of cultural mutation for Acheulean handaxe size. *PLoS ONE* 7, e48333.
- Kempe, M., Lycett, S.J., Mesoudi, A., 2014. From cultural traditions to cumulative culture: parameterizing the differences between human and nonhuman culture. *Journal of Theoretical Biology* 359, 29-36.
- Kitching, I.J., Forey, P.L., Humphries, C.J., Williams, D.M., 1998. *Cladistics: The Theory and Practice of Parsimony Analysis*, second ed. Oxford University Press, Oxford.
- Lewis, H.M., Laland, K.N., 2012. Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions of the Royal Society B* 367, 2171-2180.
- Linn, M.C., Petersen, A.C., 1986. A meta-analysis of gender differences in spatial ability, in: Hyde, J.S., Linn, M.C. (Eds.), *The Psychology of Gender: Advances through Meta-Analysis*. The Johns Hopkins University Press, Baltimore, Maryland, pp. 67-101.
- Lipscomb, D., 1998. *Basics of cladistic analysis*. <http://www.gwu.edu/~clade/faculty/lipscomb/Cladistics.pdf>
- Lycett, S.J., 2007. Why is there a lack of Mode 3 Levallois technologies in East Asia? A phylogenetic test of the Movius-Schick hypothesis. *Journal of Anthropological Archaeology* 26, 541-575.
- Lycett, S.J., 2015a. Differing patterns of material culture intergroup variation on the High Plains: quantitative analyses of parfleche characteristics vs. moccasin decoration. *American Antiquity* 80, 714-731.
- Lycett, S.J., 2015b. Cultural evolutionary approaches to artifact variation over time and space: basis, progress, and prospects. *Journal of Archaeological Science*, 56, 21-31.
- Lycett, S.J. In press. A multivariate and phylogenetic analysis of Blackfoot biographic art: Another look at the Deadmond robe. *Plains Anthropologist*.
- Lycett, S.J., von Cramon-Taubadel, N., 2015. Toward a "Quantitative Genetic" approach to lithic variation. *Journal of Archaeological Method and Theory* 22, 646-675.

- Lycett, S.J., von Cramon-Taubadel, N., Foley, R.A., 2006. A crossbeam co-ordinate caliper for the morphometric analysis of lithic nuclei: a description, test and empirical examples of application. *Journal of Archaeological Science* 33, 847-861.
- Lycett, S.J., Schillinger, K., Kempe, M., Mesoudi, A., 2015. Learning in the Acheulean: insights from experiments using handaxe form as a “model organism,” in: Mesoudi, A., Aoki, K. (Eds.), *Learning Strategies and Cultural Evolution during the Paleolithic*. Springer, New York, pp.155-166.
- Lycett, S.J., Schillinger, K., Eren, M.I., von Cramon-Taubadel, N., Mesoudi, M., 2016. Factors affecting Acheulean handaxe variation: experimental insights, microevolutionary processes, and macroevolutionary outcomes. *Quaternary International*.
<http://dx.doi.org/10.1016/j.quaint.2015.08.021>
- Martin, G., 2000. Stasis in complex artifacts, in: Ziman, J. (Ed.), *Technological Innovation as an Evolutionary Process*. Cambridge University Press, Cambridge, pp. 90-100.
- McLennan, D.A., Brooks, D.R., 2001. Phylogenetic systematics: five steps to enlightenment, in: Adrain, J.M., Edgcombe, G.D., Lieberman, B.S. (Eds.), *Fossils, Phylogeny and Form: An Analytical Approach*. Kluwer Academic / Plenum, New York, pp. 7-28.
- Mendoza-Straffon, L. (Ed.) 2016. *Cultural Phylogenetics: Concepts and Applications in Archaeology*. Springer Press, New York.
- Mesoudi, A., 2007. Using the methods of experimental social psychology to study cultural evolution. *Journal of Social, Evolutionary and Cultural Psychology* 1, 35-58.
- Mesoudi, A., 2008. Foresight in cultural evolution. *Biology and Philosophy* 23, 243-255.
- Mesoudi, A., 2011. *Cultural Evolution: How Darwinian Theory can explain Human Culture & Synthesize the Social Sciences*. Chicago University Press, Chicago.
- Mesoudi, A., O'Brien, M.J., 2008. The cultural transmission of Great Basin projectile point technology I: an experimental simulation. *American Antiquity* 73, 3-28.
- Mesoudi, A., Whiten, A., Laland, K.N., 2004. Is human cultural evolution Darwinian? Evidence reviewed from the perspective of *The Origin of Species*. *Evolution* 58, 1-11.
- Mueller, L.D., 2009. Fitness, demography, and population dynamics in laboratory experiments, in: Garland, T., Rose, M.R. (Eds.), *Experimental Evolution: Concepts, Methods, and Applications*. University of California Press, Berkeley, pp. 197-216.
- Nunn, C. L., Arnold, C., Matthews, L., Mulder, M. B. 2010. Simulating trait evolution for cross-cultural comparison. *Philosophical Transactions of the Royal Society of London B* 365, 3807-3819.
- O'Brien, M.J., Lyman, R.L., 2000. *Applying Evolutionary Archaeology: A Systematic Approach*. Kluwer Academic / Plenum, New York.

- O'Brien, M.J., Darwent, J., Lyman, R.L., 2001. Cladistics is useful for reconstructing archaeological phylogenies: Palaeoindian points from the southeastern United States. *Journal of Archaeological Science* 28, 1115-1136.
- O'Brien, M.J., Lyman, R.L., 2003. *Cladistics and Archaeology*. University of Utah Press, Salt Lake City, Utah.
- O'Brien, M.J., Boulanger, M.T., Buchanan, B., Collard, M., Lyman, R.L., Darwent, J., 2014. Innovation and cultural transmission in the American Paleolithic: phylogenetic analysis of eastern Paleoindian projectile-point classes. *Journal of Anthropological Archaeology* 34,100-119.
- Okumura, M., Araujo, A.G.M., 2014. Long-term cultural stability in hunter-gatherers: a case study using traditional and geometric morphometric analysis of lithic stemmed bifacial points from Southern Brazil. *Journal of Archaeological Science* 45, 59-71.
- Prentiss, A.M., Walsh, M.J., Foor, T.A., Barnett, K.D., 2015. Cultural Macroevolution among high latitude hunter-gatherers: a phylogenetic study of the Arctic small tool tradition. *Journal of Archaeological Science* 59, 64-79.
- Riede, F., 2008. Maglemosian memes: technological ontogeny, craft traditions, and the evolution of Northern European barbed points, in: O'Brien, M. (Ed.), *Cultural Transmission and Archaeology: Issues and Case Studies*. Society for American Archaeology Press, Washington, D.C., pp.178-189.
- Riede, F. 2009. Tangled Trees: Modeling Material Culture Evolution as Host-Parasite Cospeciation, in Shennan, S. (Ed.) *Pattern and Process in Cultural Evolution*. University of California Press, Berkeley, pp.85-98.
- Robert, M., Chevrier, E., 2003. Does men's advantage in mental rotation persist when real three-dimensional objects are either felt or seen? *Memory & Cognition* 31, 1136-1145.
- Rohlf, F.R., 2010. TpsDig2 v2.16. Department of Ecology and Evolution, State University of New York, Stony Brook, New York. <http://life.bio.sunysb.edu/morph/>
- Schiffers, M.B., Skibo, J., B., 1987. Theory and experiment in the study of technological change. *Current Anthropology*, 595-622.
- Schillinger, K., Mesoudi, A., Lycett, S.J., 2014a. Copying error and the cultural evolution of "additive" versus "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, 446-455.

- Schillinger, K., Mesoudi, A., Lycett, S.J., 2016. Differences in manufacturing traditions and assemblage-level patterns: the origins of cultural differences in archaeological data. *Journal of Archaeological Method and Theory*, in press.
- Schuh, R.T., Brower, A.V.Z., 2009. *Biological Systematics: Principles and Applications*. Ithaca, Cornell University Press.
- Shennan, S., 2000. Population, culture history, and the dynamics of culture change. *Current Anthropology* 41, 811-835.
- Shennan, S.J., 2011. Descent with modification and the archaeological record. *Philosophical Transactions of the Royal Society B* 366, 1070-1079.
- Sokal, R.R., Rohlf, F.J., 1995. *Biometry*. W.H. Freeman & Co, New York.
- Spencer, M., Davidson, E.A., Barbrook, A.C., Howe, C.J., 2004. Phylogenetics of artificial manuscripts. *Journal of Theoretical Biology* 227, 503-511.
- Swofford, D.L., Sullivan, J., 2009. Phylogeny inference based on parsimony and other methods, in: Lemey, P., Salemi, M., Vandamme, A.-M. (Eds.), *The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing*. Cambridge University Press, Cambridge, pp. 267-312.
- Tehrani, J.J., 2011. Patterns of evolution in Iranian tribal textiles. *Evolution: Education and Outreach* 4, 390-396.
- Tehrani, J.J., Collard, M., Shennan, S.J., 2010. The cophylogeny of populations and cultures: reconstructing the evolution of Iranian tribal craft traditions using trees and jungles. *Philosophical Transactions of the Royal Society B* 365, 3865-3874.
- Tomasello, M., 1999. *The Cultural Origins of Human Cognition*. Harvard University Press, Cambridge, MA.
- Voyer, D., Voyer, S., Bryden, M.P., 1995. Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin* 117, 250-270.
- Whiten, A., Horner, V., Litchfield, C.A., Marshall-Pescini, S., 2004. How do apes ape? *Learning & Behavior* 32, 36-52.
- Wynn, T., Tierson, F.D., Palmer, C.T., 1996. Evolution of sex differences in spatial cognition. *Yearbook of Physical Anthropology* 39, 11-42.

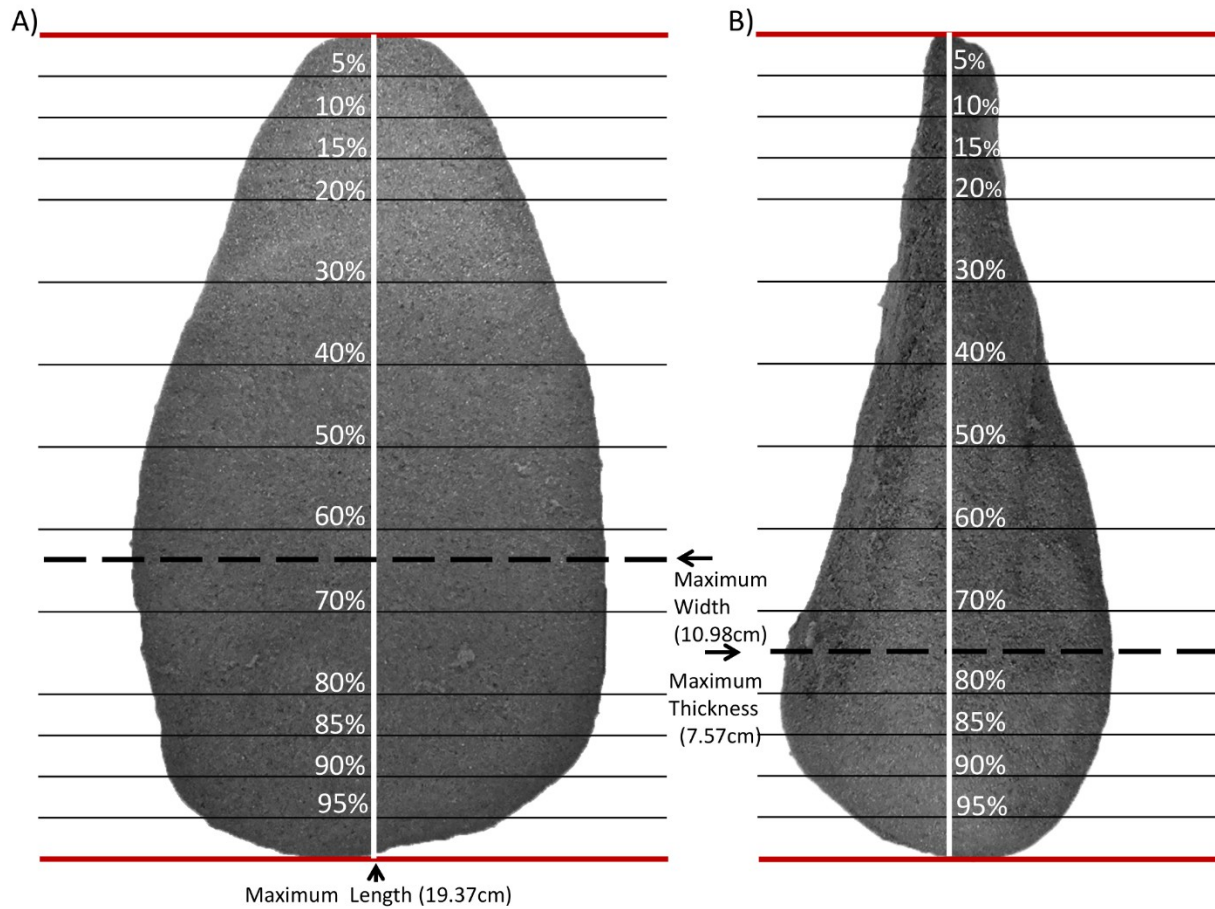


Figure 1. Plan-view (a) and profile-view (b) of the target “model” handaxe given to participants in order to copy (maximum length, width and thickness dimensions indicated). Percentage points indicate points at which bilateral width and thickness measurements were taken in order to provide morphometric comparative data.

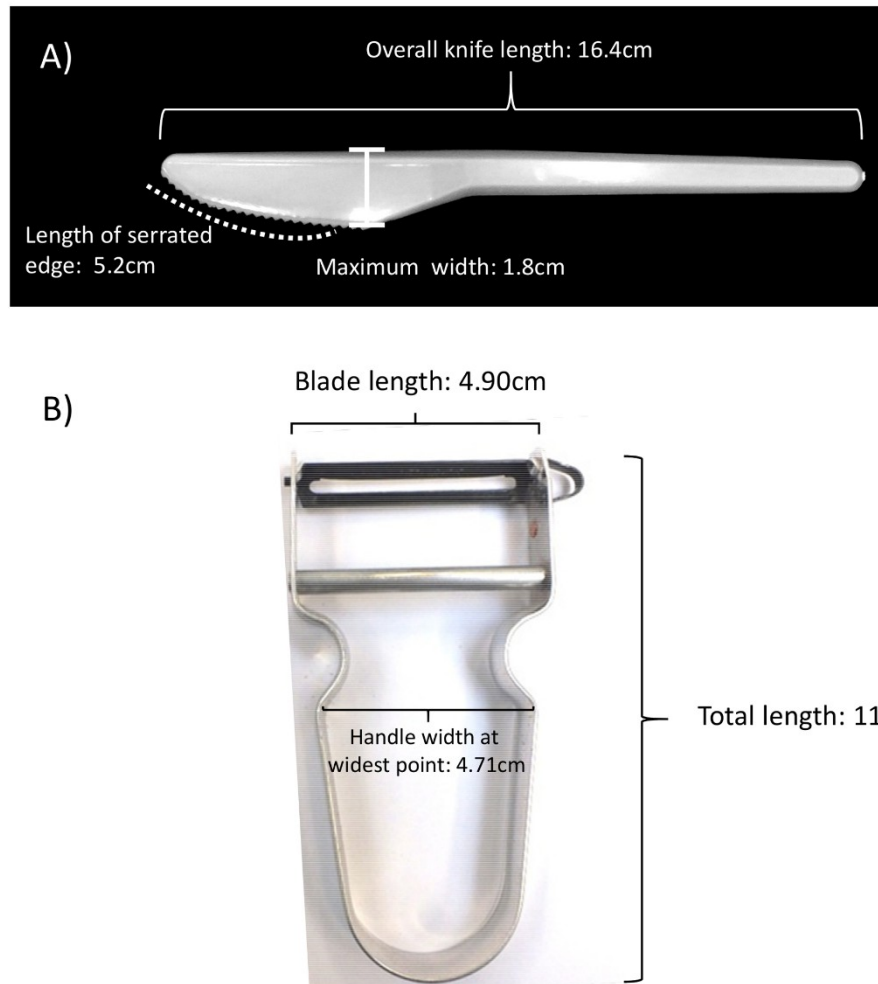


Figure 2. Example of plastic knife (a) and metallic peeler (b) provided to experimental participants in order to shape their foam artifact copies.

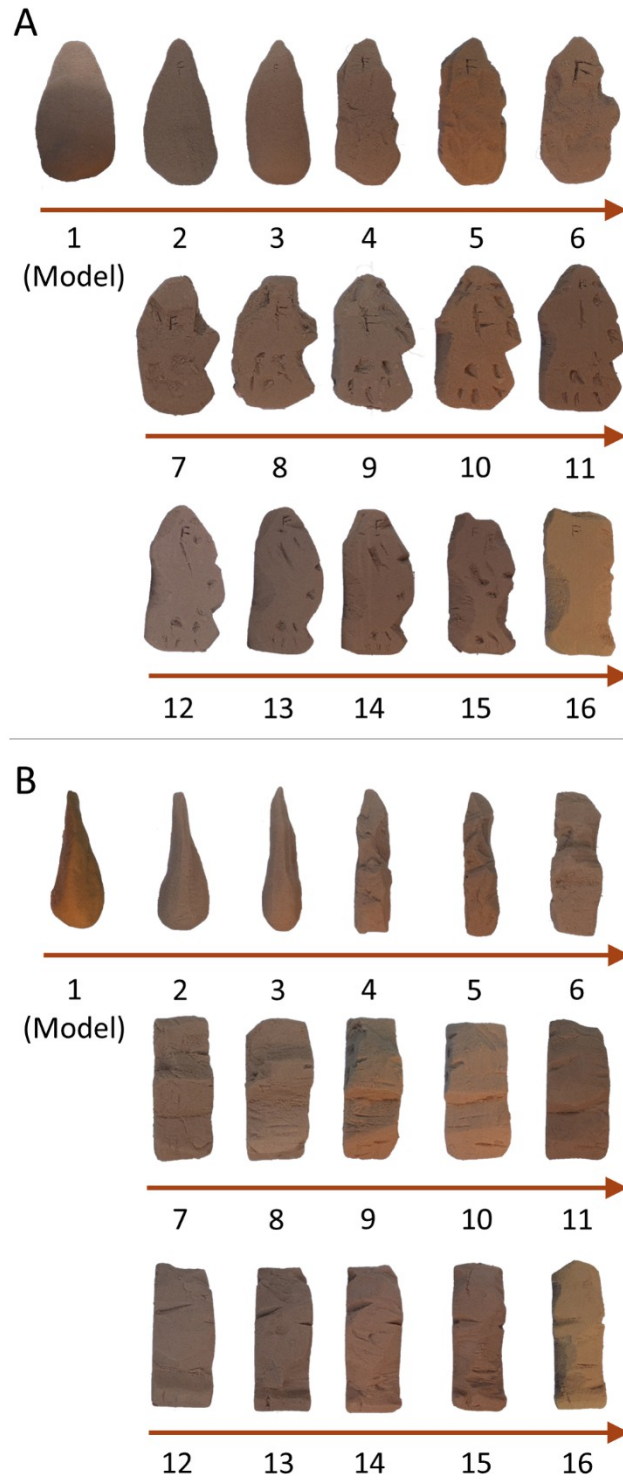


Figure 3. Plastic-knife transmission chain in (A) plan view and (B) profile views.

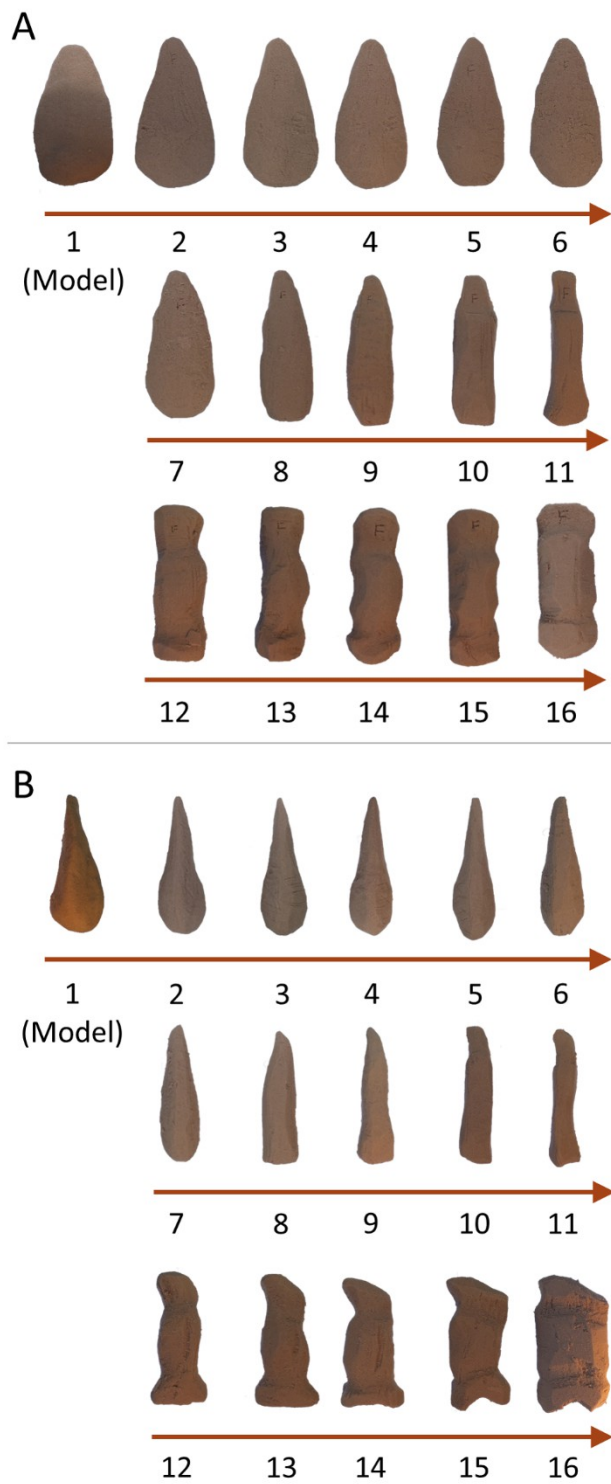


Figure 4. Metallic-peeler transmission chain in (a) plan view and (b) profile views.

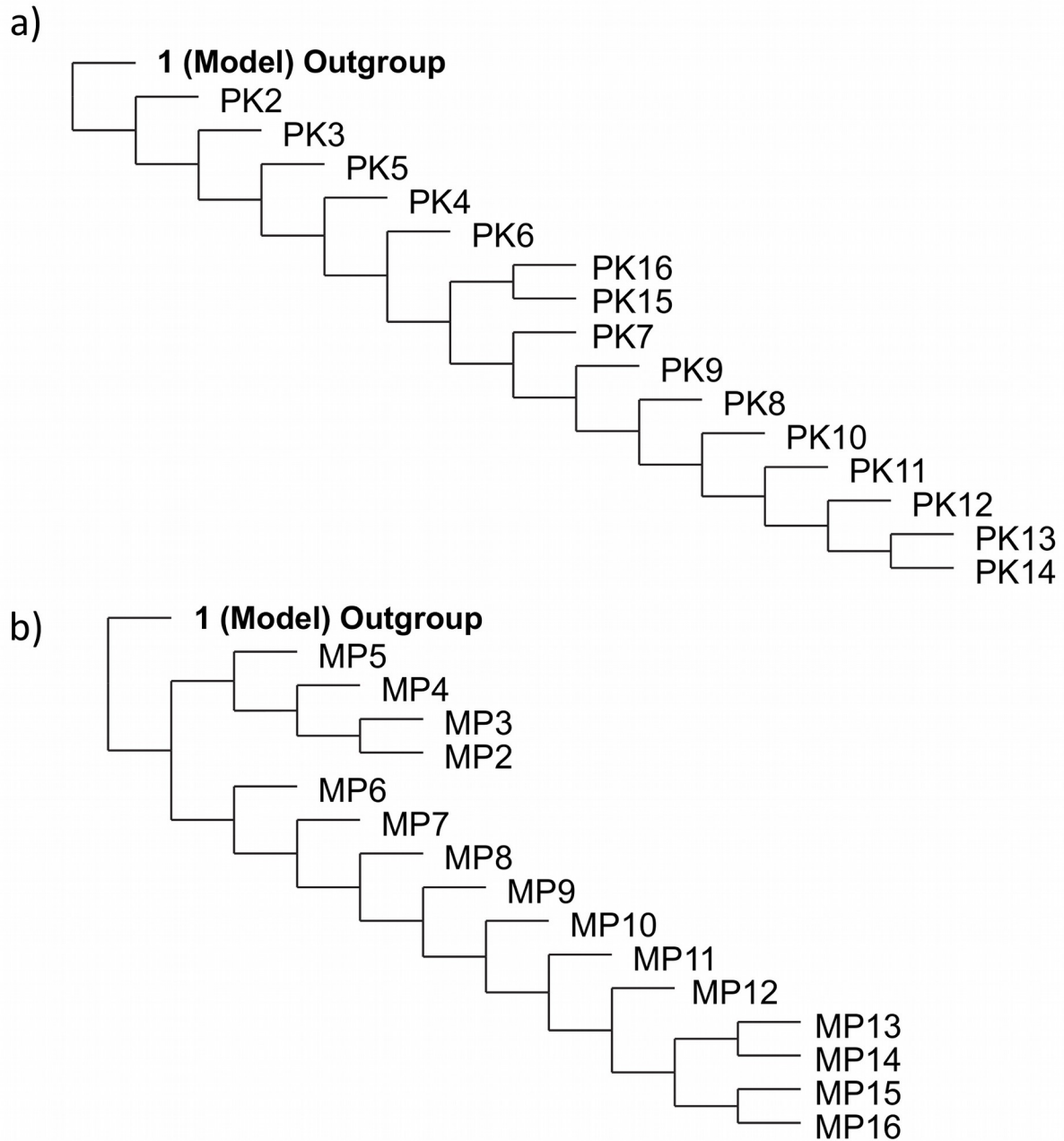


Figure 5. (a) Single most-parsimonious tree produced for plastic-knife chain. Tree length (TBR score) = 49.13, Retention Index = 0.66; (b) Single most-parsimonious tree produced for the metallic-peeler chain. Tree length (TBR score) = 48.25, Retention Index = 0.81.

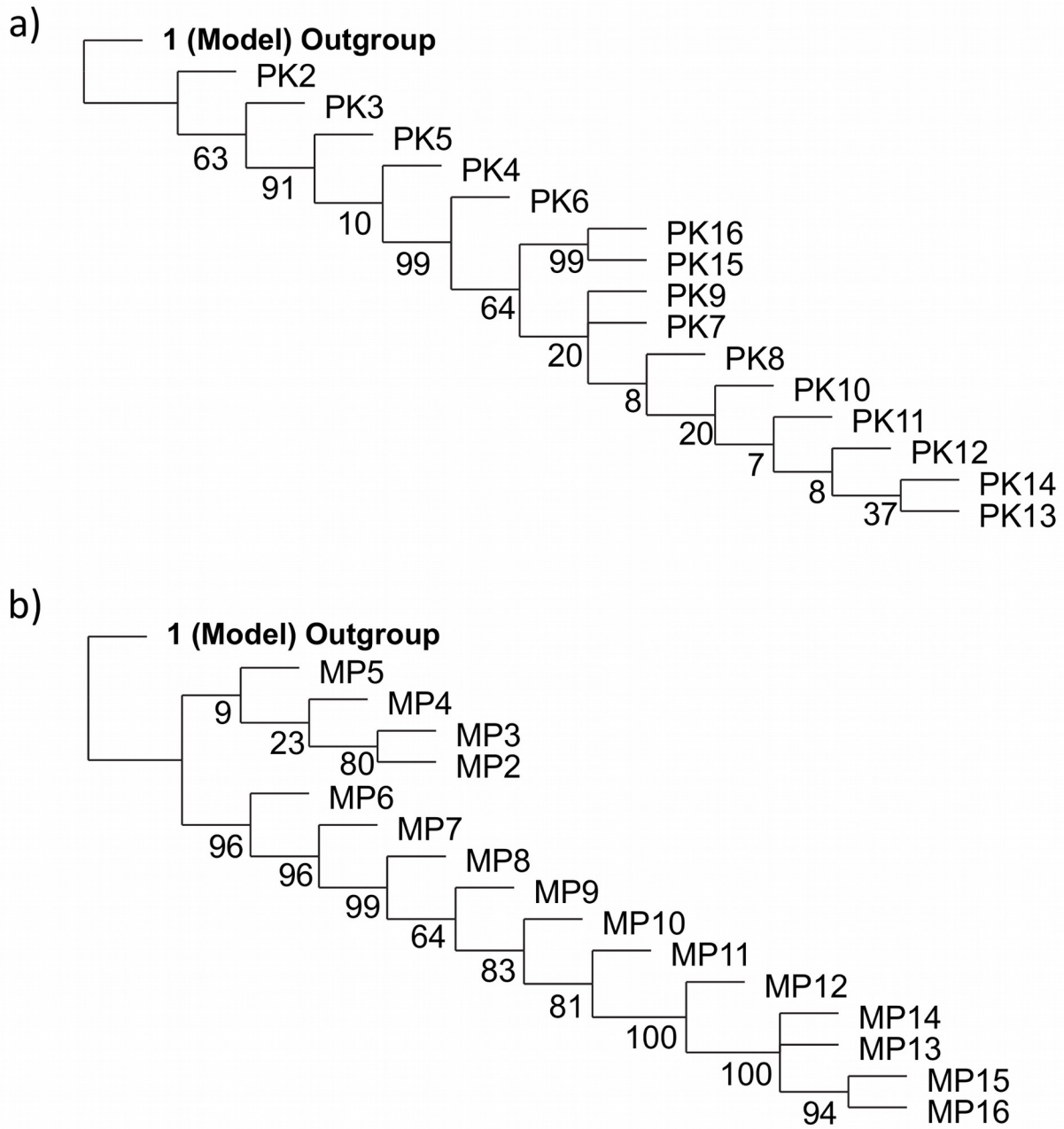


Figure 6. (a) Results of bootstrap analyses for the plastic-knife transmission chain data; (b) Results of bootstrap analyses for the metallic-peeler transmission chain data.

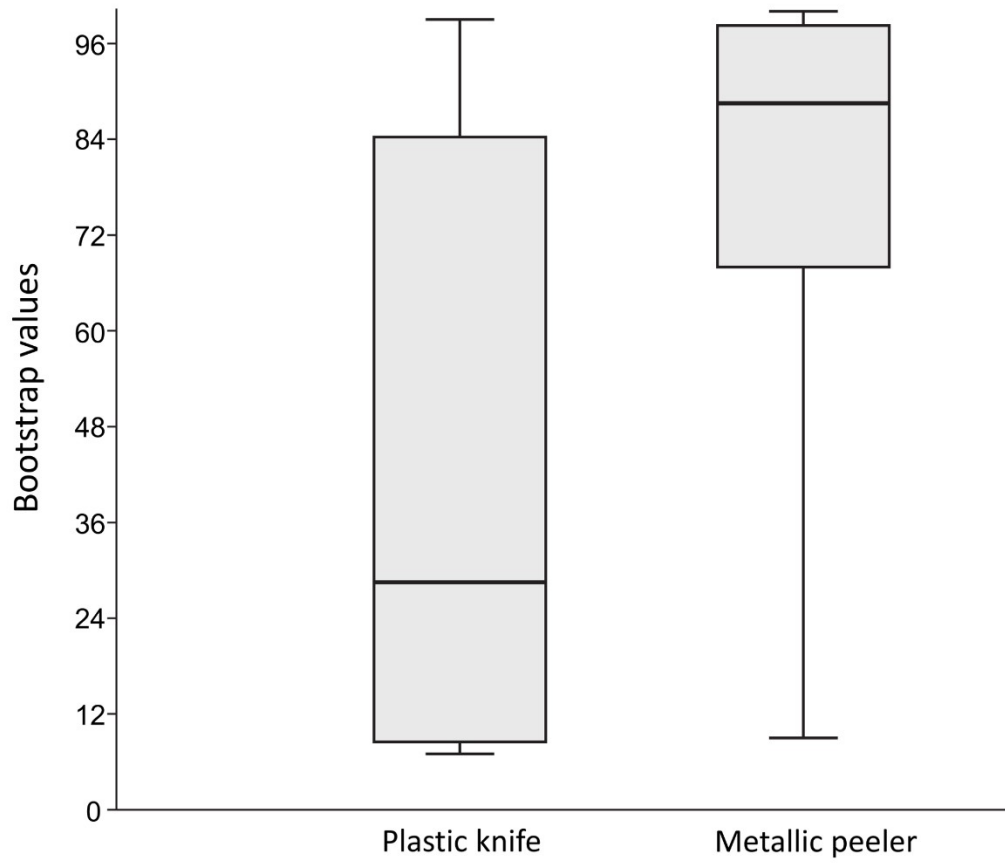


Figure 7. Box plots showing differences in bootstrap values for the plastic-knife and metallic-peeler experimental conditions. In each case, boxes indicate 25–75 percent quartiles and horizontal bars indicate median values.