Tales of tools and trees

Phylogenetic analysis and explanation in evolutionary archaeology

Wybo Houkes

Philosophy & Ethics of Technology

Eindhoven University of Technology

IPO 1.06; Den Dolech 2

NL-5600 MB Eindhoven

The Netherlands

w.n.houkes@tue.nl

Abstract

In this paper, I study the application of phylogenetic analysis in evolutionary archaeology. I show how transfer of this apparently general analytic tool is affected by salient differences in disciplinary context. One is that archaeologists, unlike many biologists, do not regard cladistics as a tool for classification, but are primarily interested in explanation. The other is that explanation is traditionally sought in terms of individual-level rather than population-level mechanisms. The latter disciplinary difference creates an ambiguity in the application and interpretation of phylogenetic analyses. Moreover, I argue that, while archaeologists have claimed that "cladistics is useful for reconstructing artefact phylogenies" (O'Brien et al. 2001), these reconstructions only contribute minimally to the explanatory research agenda of evolutionary archaeology.

Tales of Tools and Trees

Phylogenetic analysis and explanation in evolutionary archaeology

1. Darwinizing culture: global and local revolutions

Evolutionary theory is outgrowing its natural habitat. Many researchers outside biology frame their questions and results in evolutionary terms, and propose counterparts to mechanisms and entities that are central to our understanding of the organic world. Thus, historians and economists discuss how introducing a genotype-phenotype distinction helps to describe technological change (e.g., Ziman 2000) and engineers claim that solutions of design problems may evolve by natural selection (e.g., Bentley 1999).

This 'second Darwinian revolution' has not escaped philosophical scrutiny. Philosophers of science have examined, for instance, which explanatory role could conceivably be played by evolutionary mechanisms with respect to social processes (Sober 1991) or the spread of religious ideas (Sterelny 2006). The question that drives most of this work is whether there is, given the existence and success of alternative accounts, any explanatory work done by or left to do for evolutionary accounts of culture and technology. The critical focus is on general theories of cultural evolution, such as dual-inheritance theory or memetics (see Laland and Brown 2002 for a useful survey).

Not all extensions of Darwinism takes this global form. Much consists of selective applications, to specific disciplinary problems, of tools or techniques taken from evolutionary biology. Research papers in evolutionary economics, engineering and archaeology rarely mention general frameworks such as dual-inheritance theory and only occasionally discuss the possibilities of defining suitably general evolutionary concepts. Instead, the results reported are gained by applying specific tools and techniques. Economists have, for example, adopted Lotka-Volterra models to study how technological change affects competition between firms (e.g., Saviotti 1995); and archaeologists and anthropologists have recently started using phylogenetic analysis to reconstruct tool traditions. The latter efforts will be the subject of this paper.

Applications of evolutionary tools outside biology raise methodological problems which have been largely ignored by philosophers of science. Critical responses are instead brought forward locally, by researchers within the discipline who favour a non-evolutionary approach. One objection is found throughout the literature, although the details vary with the evolutionary tool being questioned. Critics maintain that culture and technology essentially involve human beliefs and intentions. From this, they conclude that evolutionary tools, such as replicator dynamics or phylogenetic reconstruction, structurally misrepresent the phenomena. A standard response is to point out that, although methodological problems may arise in applying evolutionary tools to cultural phenomena, the problems are largely the same as those in biology. Cultural evolution is presented as 'business-as-usual' from a biological perspective: not less difficult, but also not less appropriate to the domain.

In this paper, I develop another perspective on the application of evolutionary tools to Darwinize culture. Instead of focussing on structural (dis-)similarity, I examine how the change of disciplinary context affects the relevance of evolutionary tools. This leads to a detailed appreciation of the difficulties in transferring one item in the evolutionary toolkit from the organic to the artificial world. I first describe phylogenetic techniques and their application in archaeology (section 2). Although phylogenetic reconstructions are strictly speaking neutral tools to analyse similarity relations between items, I go on to show that the relevance conditions set by the disciplinary context differentiate applications of these tools in archaeology from those in biology. In archaeology, these reconstructions are supposed to contribute to a primarily explanatory project (section 3), which conceals an important ambiguity regarding the type of explanation sought – and to which, given our present knowledge of (evolutionary) mechanisms underlying design and use practices, the construction of tool trees contributes little (section 4). I also offer suggestions regarding a productive research agenda and the role of phylogenetic reconstructions therein.

2. Trees of tools: How phylogenetics came to archaeology

Before describing the use of phylogenetic analyses in archaeology, some words about the technique in general. One of Darwin's central insights was that of the common ancestry of (all of) life. All biological species are related by sharing, somewhere in their part, common ancestral species. The historical relations within any set of species can therefore be represented by a branching, tree-like structure, from the ancestor they all share at the root to the topmost branching points at which one finds (more recent) ancestors shared by only two species. A sample phylogenetic tree, for five species, is given in Figure 1. All species share a common ancestor (F^*) and species A and B share a more recent common ancestor (G^*) than species A and D (H^*).
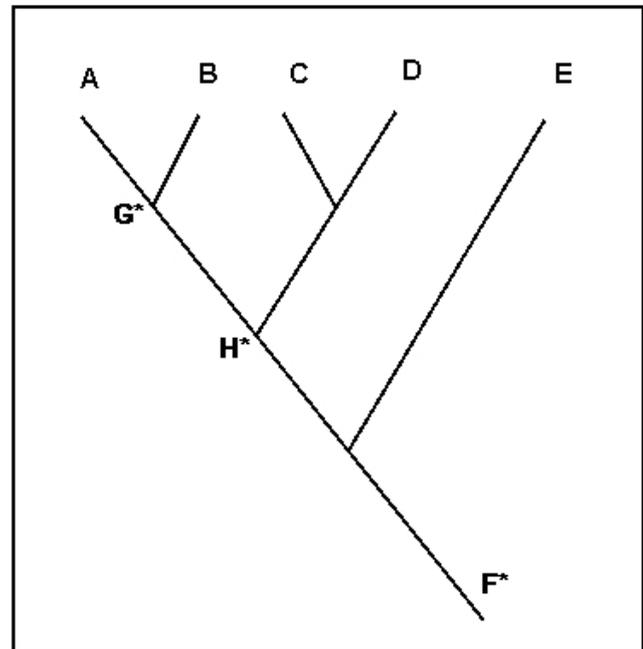

Figure 1. A sample phylogenetic tree

Various techniques exist for reconstructing phylogenetic trees. Most are based on morphological characteristics of the species studied,[1] typically called "states" of "characters". States can be discrete or continuous; in practice, continuous states are often turned into discrete ones by introducing cut-off points. Sample characters are feeding habits (with states "carnivorous", "herbivorous" and "omnivorous") or adult wingspan (with continuous states and a cut-off point at one metre). Starting from these characteristics, a number of methods have been developed to infer phylogenies, most of which are implemented in software packages. A popular choice, the most popular among archaeologists, is the "maximum parsimony" method, implemented in PAUP (Swofford 2002). This reconstructs the tree by finding the minimum number of evolutionary events that can account for the observed characteristics – roughly

---

[1] In biology, phylogenetic reconstruction is increasingly based on molecular rather than morphological data. Some authors (e.g., Ridley 2004, Ch.16) therefore distinguish 'molecular phylogenetics' from (morphology-based) 'cladistics'. Evolutionary archaeologists use cladistics more broadly, to denote any method of phylogenetic reconstruction, and exclusively use morphological data.

speaking, it minimises the total number of character-state changes in the tree. Another choice is that of an outgroup, a species that is phylogenetically related to, but not included among, the set of species involved. This choice converts a network of similarities into a rooted tree, and is non-trivial because it requires prior knowledge of some phylogenetic relations.

On the basis of these choices, phylogenetic analysis not only determines relations of common ancestry, but also identifies the character states of hypothetical common ancestors; in fact, the former is achieved by the latter. Thus, the result looks conspicuously like a family tree in which portraits are proposed for unknown ancestors. On the basis of the analysis, similarities between items can be distinguished into (at least) two classes: *homologies*, character states that are shared between items and their common ancestor; and *homoplasies*, states that are shared between items, but not with their most recent common ancestor.[2] To give an example: the wings of parrots and eagles are homologous, whereas parrot wings and bat wings are homoplasious.

It is difficult to assess independently the outcome of phylogenetic analyses: data are typically scarce, and relations of ancestry are not directly observable. One thing that distinguishes parsimony-based cladograms from systematically generated historical narratives is that there are two measures of quality: the 'consistency' and 'retention' indexes of the cladogram. These numbers measure how well the data fit the branching structure 'imposed' on it in phylogenetic analysis;[3] the measured quantity is sometimes described as the strength of the phylogenetic signal in the data set. The more homoplasies are needed to reconstruct the relations of common ancestry between the items, the lower the signal and the indexes.

Parsimony methods for phylogenetic reconstructions are well-established in systematic biology, and have also been applied in other disciplines, like linguistics and archaeology. In archaeology, the application takes the shape of reconstructing 'tool traditions', lineages of artefacts that are roughly functionally equivalent.

---

[2] Whether phylogenetic analyses provide independent evidence for the homology-analogy distinction, let alone sufficient evidence, is a thorny issue. It is left aside in this paper, because it intuitively affects applications in biology as much as those in archaeology.

[3] The consistency index (CI) is the ratio between the minimum possible number of character changes and the number of changes on the resulting tree. The retention index (RI) is a slightly more complicated ratio that does not depend on the size of the data set. The differences between these indexes are irrelevant for the purpose of this paper.

Although more qualitative reconstructions were suggested earlier (see, e.g., Basalla 1988), a rigorous application of parsimony-based phylogenetic analysis is of recent date (O'Brien et al. 2001). It involves the choices described above: characters and states, outgroup and analytic method (parsimony).[4] The result, of which Figure 2 represents a small section, is a cladogram for paleoindian projectile points found in the Southeastern Basin of the United States. The full cladogram shows the lineage of seventeen classes of these points, including their common ancestors (on the branching points). The boxes on the lines represent character-state changes, differentiated by type: open boxes are unique changes in the ancestral state; shaded boxes represent state changes that occur more than once; and half-shaded boxes are reversals to the ancestral state.
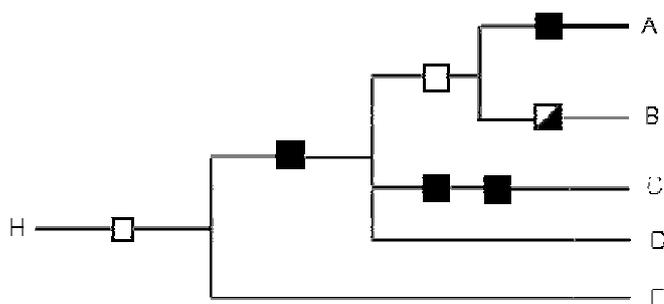

Figure 2. A cladogram for paleoindian projectile-points (after O'Brien et al. 2001: 1132)

Recently, the phylogenies of other tool traditions and artefact lineages, including Copador ceramics and Turkmen textiles, have been reconstructed.[5] This adoption of the technique has also triggered criticisms from fellow archaeologists and anthropologists. Their objections fit the template described above. In particular, it has been pointed out that the tree of technology cannot have the neat branching structure of the tree of life. Rather, it should show many intersections of lineages due to blending or reticulation of cultural/tool traditions. One cause of this blending would be "pervasive theft of ideas", as one critic, Niles Eldredge (introduction to Lipo et al. 2006: xvi), has put it: human beings, including artisans, imitate each other easily and

---

[4] Another choice is to focus on the characters of *artefacts* instead of those of, e.g., use practices or larger cultural units. This is controversial, also among evolutionary archaeologists.
[5] Mace et al. (2005) and Lipo et al.(2006) provide (references to) more examples of phylogenetic reconstructions in archaeology.

often. Applying cladistics would therefore give a distorted or incomplete representation of the history of culture and technology.

Advocates of artefact phylogenetics defend their approach with an "as-in-nature" argument; in this case, a comparison of the quality of phylogenetic analyses in systematic biology and those in anthropology and archaeology (Collard et al. 2006). The consistency and retention indexes for 'cultural' cladograms lie in the same range as those in nature, meaning that cladistic techniques are as appropriate to cultural phenomena as they are to biological species: cladistics misrepresents the tree of technology as much (or little) as it misrepresents the tree of life. Since archaeologists face the same methodological choices as systematic biologists, there is structural and methodological similarity between cultural and natural evolution.

A remarkable illustration of the stalemate to which this focus on structural (dis-) similarities has led is the application of phylogenetic techniques to two types of musical instruments (Tëmkin and Eldredge 2007). This application is intended to show how a branching structure misrepresents the actual history of these instruments, which has a much more reticulated structure. Advocates of artefact phylogenetics have welcomed this intended reductio as a *result*; they emphasise that the retention index of, for instance, the Baltic psaltery cladogram lies just below the 'cultural average' – which is equal to the biological average. A non-partisan presentation mentions spearheads, textiles and musical instruments on a par, as items for which "phylogenetic trees have been constructed" (Venditti and Pagel 2008).

3. Clades in classification and explanation

Suppose, for the sake of the argument, that tool trees are indeed structurally similar to the tree of life, and in reconstructing the former one faces the same methodological choices and dilemmas as in reconstructing the latter. This equivalence would leave room for another type of criticism. To put it metaphorically: the consistency and retention 'scores' of trees of tools and of life may be the same – but are archaeologists and biologists playing the same game in applying cladistics? In this section and the next, I show how the application of cladistic techniques in archaeology is affected by its disciplinary context, which is different from that in biology.

One difference concerns the broader aim of phylogenetic analyses – the reason why researchers in a discipline are interested in trees. In biology, the relevance of

cladistics lies in large part in their use for systematics, i.e., for the correct classification of living items into groups. Phylogenetic analyses contribute directly to determining the historical relations between these groups, which are of obvious importance in their classification and, arguably, in their nomenclature. Indeed, "cladistics" not only indicates a popular type of (parsimony-based) phylogenetic analysis, but also an approach to the classification of living items – showing a close association between the analysis and its expected relevance. Consequently, many applications of and discussions regarding the analyses concern revisions of traditional classifications (e.g., Mayr 1981; Hull 1988). One example is the classification of birds and various dinosaur species. Cladistics has been used to determine historical relations between these species, which has contributed significantly to the ongoing debate about the evolution of birds from dinosaur-like, lizard-like or other ancestral species.

Phylogenetic analyses are not exclusively applied for classification. As the origin-of-birds example shows, issues of classification are connected to explanations, e.g., of the evolutionary "events" identified in the phylogenetic analysis. Another application is in phylogenetic comparative methods (e.g., Harvey and Pagel 1991), which are used to test correlations for independence of data – and thus to evaluate hypotheses about the adaptation of species to their environment or the co-evolution of traits.

However, phylogenetic analyses would be relevant to biology even if they would contribute only minimally to explanatory projects. Archaeologists are, by contrast, uninterested in classification. Most regard any system of distinguishing artefact kinds – or, more generally, cultural kinds – as "conventional" or interest-dependent. As an influential textbook puts it: "Types, assemblages and cultures are all artificial constructs designed to put order into disordered evidence (...) [D]ifferent classifications are needed for the different kinds of questions we want to ask" (Renfrew and Bahn 2004: 118). This conventionalism has been widely accepted in archaeology following a debate in the 1950s about whether artefact kinds are discovered or constructed.[6]

---

[6] O'Brien and Lyman (2000, pp. 207-213) summarize this so-called 'Ford-Spaulding' debate.

This view may be in need of critical scrutiny,[7] but it is taken as a given about the disciplinary context here. It entails that archaeologists cannot find phylogenetic analyses relevant because it could contribute to classification: classification is supposed to precede analysis, not be an outcome of it. Instead, the main relevance of phylogenetic reconstruction may lie in its contribution to explanations. This would be in line with the overarching goal of the (sub-) field. A textbook describes the challenge for evolutionary archaeology as to "explain [processes of cultural change] more coherently or persuasively than hitherto" (Renfrew and Bahn 2004, p.481). This is underwritten by evolutionary archaeologists themselves:

> "The Darwinian mechanisms of selection and transmission, when incorporated into an explanatory theory, provide precisely what culture historians were looking for: the tools to begin explaining cultural lineages—that is, to answering why-type questions." (O'Brien and Lyman 2002: 35)

The concern with explanation is also evident in descriptions of the results of artefact cladistics, e.g., in the following quote:

> "We use cladistics because of its unique ability ... to lay out sequences of character-state changes. (...) The ultimate goal is to explain why side notching was preferred over other hafting techniques" (Darwent and O'Brien 2006: 185)

The association with classification is therefore exchanged for a focus on explanation. This raises the question, however, how phylogenetic reconstructions contribute to explanations – and to which explanations they contribute.

---

[7] E.g., in light of the recent debate about the realism of artefact kinds (see, e.g., the contributions by Elder and Thomasson to Laurence and Margolis 2007)

4. Two tales of tools, but not with trees

In archaeology, advocates of evolutionary tools should prove the merits of their tools by answering outstanding why-questions. For artefact lineages, this explanatory project is thought of as consisting of two steps: first, the cladogram is constructed; next comes the archaeologists "truly grueling ... job" (O'Brien and Lyman 2000: 20) of explaining the shape of the cladogram. This process of first constructing and then explaining or interpreting a cladogram is echoed in more places, including the earlier quote about explaining cultural lineages through evolutionary mechanisms.

In *Reconstructing the Past*, Elliott Sober analyses a particular kind of explanation that apparently underwrites this two-step process, but also shows some general difficulties in taking the second step. Sober (1988, p.133) notes that what he calls "genealogical explanations of observed character distributions" – a particular kind of historical explanation of varieties of characteristics in a data sample – require three ingredients, two of which are important in the present context: a genealogical hypothesis about character states; and mechanisms that explain changes of character states.[8] This shows how phylogenetic reconstructions can contribute to explanations, since cladograms represent a genealogical hypothesis about common ancestry. It also shows that cladograms do not provide explanations without hypotheses about mechanisms for state changes.

Archaeologists are aware that explanations cannot be read off directly from cladograms and that finding explanations is difficult: the full description of the second step in the source quoted above is "a truly grueling *and often fruitless* job" (O'Brien and Lyman 2000: 20; emphasis added). Still, the need for mechanisms raises two questions about the explanatory project envisaged. First, which mechanisms are available, in terms of background knowledge, to explain the structure of artefact lineages? Secondly, given the need for and perhaps choice of mechanisms, is constructing the cladogram independent from explaining it?

In this section, I argue for three critical claims: that archaeologists make different and implicit choices regarding the level of explanation; that these choices affect phylogenetic reconstructions, so that the explanatory step cannot be neatly separated from the reconstructive step; and that, on either of the two choices regarding

---

[8] The third ingredient are "interior states", a claim about character states of common ancestors and (lack of) changes in these states.

mechanism, phylogenetic analyses contribute little or nothing to the explanatory project of (evolutionary) archaeology.

First, the ambiguity. Evolutionary archaeologists may and do seek explanatory mechanisms at two different levels. On the one hand, they can appeal to broad knowledge bases from everyday life and more systematic inquiries, including anthropological case studies, experimental archaeology and innovation studies. These sources provide information about the transmission and change of technological knowledge and artefacts, including training and selective communication; imitation and theft of ideas; independent invention and guided variation; and modularity of design problems and solutions. In most real-life cases, combinations of these mechanisms will operate, creating an overwhelming variety of possible explanatory mechanisms. These mechanisms focus on decisions made by individual users and designers of artefacts, against a social background. Moreover, the artefacts are mostly individual items or kinds with few members – archaeologists do not describe an era of mass-production.[9] Pitched at this level, the efforts of archaeologists would resemble the application of cladistic techniques to the genealogy of copied texts such as *The Canterbury Tales* or chain letters (e.g., Spencer et al. 2004). The resulting textual-tradition trees feature *individual* texts and may be interpreted in terms of explanatory mechanisms such as transcription errors, made by individual scribes, and "deliberate contamination" from other texts than the copied one.

Evolutionary theory offers another set of mechanisms, such as natural selection and drift. These mechanisms, which operate on the population level, should be distinguished from various "selective forces" operating on the individual level.[10] Explanations in archaeology typically do not involve this population level, and it is not self-evident that populations of projectile points and other technical artefacts evolve by natural selection (more on this below). However, only if they are pitched at this level, the explanations given by archaeologists would resemble those given by evolutionary biologists.

So which choice is made? The explanations given by evolutionary archaeologists, despite their self-chosen name, often reveal an overriding interest in

---

[9] Historical relations between mass-produced items may also be susceptible to an individual (artefact-type) level of explanation, since tokens of the type are, for most practical purposes, identical. This raises the question – not addressed in this paper – whether artefact kinds are sufficiently like biological species to reconstruct their historical relations by the same methods.
[10] Here, I follow Lewens (2004), pp.22-28

the individual level. For instance, in interpreting the projectile-point cladogram above, researchers note that: "*Unfortunately*, the results of cladistics cannot determine whether *a character change* represents a functional change in projectile-point design, nor do we have informants who can tell us *why they changed a particular feature*. However, the engineering properties of different characters can give us some insight" (Darwent and O'Brien 2006: 200; emphasis added). This shows that the information sought concerns choices made by individual projectile-point designers, much like the transcription errors of individual scribes. The authors note that interviews cannot reveal this information and go on to invoke engineering properties, e.g., material constraints in the design process. This invokes background knowledge that other archaeologists will immediately recognize – just like the individual-level interest.

This interest also affects the construction of the cladogram, which leads to the second critical claim. The diagram shown in Figure 2 features seventeen *classes* of points, which appears to put the diagram at a population level. However, these classes have only a few members (four or more) and were chosen out of five hundred classes from a sample of around six hundred projectile points. Most classes therefore have only one member.[11] Moreover, multiple-member kinds were not selected because of an interest in artefact populations, but to decrease the idiosyncrasies caused by small data samples. Therefore, the cladogram in Figure 2 is naturally interpreted as the history of individual spearheads and their makers and users: the boxes may be regarded as the results of instances of (combinations of) individual-level mechanisms such as plagiarism or guided variation under strong material constraints. Thus, it would seem that the individual-level interest has implicitly guided the choice of artefact 'kinds' that enter the construction of the cladogram. The two steps of the explanatory project only seem separable as long as the choice of explanation-level is left implicit.

This may suggest that phylogenetic reconstructions of tool trees would best be pitched explicitly at the level of individual items and relevant selective forces. After all, this would fit the interest and background knowledge of archaeologists; and

---

[11] This may be interpreted as a result of scarcity of data, a problem archaeologists share with paleontologists. This interpretation regards items as representatives (prototypes) of a larger population rather than as items of individual interest. I have found no evidence in the literature that archaeologists chose this interpretation; if they do, their implicit explanatory pitch is at the population level, discussed later in the main text.

cladistic techniques have proven their worth in reconstructions of manuscript traditions, which operate on the same level.

One consequence of this choice is that the resulting reconstruction or explanation is not evolutionary, because it is not populational. This does not discredit the explanations – many respectable historical explanations are not populational or Darwinist – but it does make it misleading to present the explanatory project as *evolutionary* archaeology.

A more principled problem is that an overwhelming variety of mechanisms operates on the level of individual artefacts, many or all of which involve human intentions. Even setting aside the specific complexities of validating intentional explanations, explanatory hypotheses are vastly underdetermined by the available evidence. The problem with archaeological explanations is not that few hypotheses fit the (scant) facts, but that there is a "superabundance" (Renfrew and Bahn 2004, p.469), with no principled way to prefer one explanation over the other.

Do cladograms contribute to solving this difficulty? One might think they do. After all, they make a principled distinction between homologies and homoplasies. This means that some similarities are unlikely to be the result of independent invention, like homologies in biology are not cases of convergent evolution. Yet an abundance of explanatory mechanisms is left: homologies could be explained by strong material constraints on design, fixed (but false) technological knowledge about these constraints, strict master-apprentice relations, unchanging user preferences, to name but a few; likewise, homoplasies can be the result of – among other things – prestige or conformity bias among cultures, independent invention, and dependence between traits. Moreover, for any change in character state, one can appeal to changes in either design or use practices.

One could counter that excluding one of many mechanisms would improve the explanatory situation. However, cladistic techniques might also overlook or downplay the importance of some mechanisms, causing loss of information. Here, on the individual level, the structural-misrepresentation objection appears to have some bite after all. Background knowledge tells us that, for instance, theft of ideas and modularity of design problems are significant forces on the individual level. The resulting similarities may, in the case of artefacts, jump over generations as old solutions to partial problems are borrowed. 'Ordinary' theft may appear on the cladistic radar, so to speak, as a homoplasy. Intergenerational theft and borrowing,

however, are unlikely to do so – and the more the situation resembles one of multiple ancestry, the less likely the situation is to be represented accurately. The problem is not that there is (too much or any) blending or reticulation, but that one might not see all there is to explain "on the cladistic radar". It is an open question whether this limited misrepresentation objection holds true. Tests with small, experimentally controlled tool traditions need to show to what extent the strong modularity and material constraints inherent in many technologies undermine the representational accuracy of cladograms.[12] Without successful tests, the possibility of neglecting significant individual-level mechanisms in phylogenetic reconstructions should give one pause in applying these reconstructions for explanatory purposes.

One might therefore seek strength in numbers, i.e., pitch explanations at the population level. Indeed, the promise made by evolutionary archaeologists (quoted in section 3) seems to be that the abundance of individual-level mechanisms may be *replaced* with the population-level mechanisms of evolutionary theory.

To move to the population level, the items that are subjected to phylogenetic analysis should be artefact *types* rather than individual tokens. Suppose that this choice, non-standard in archaeology, could be made in some uncontroversial manner. Even then, given the present state of our knowledge, the phylogenetic reconstruction would contribute little to population-level archaeological explanations. The reason is that one cannot simply assume or postulate that populations of material artefacts evolve by natural selection. The seminal work on dual-inheritance models of human cultural evolution by Boyd and Richerson (1985) shows that some cultural items, such as child-bearing practices, may evolve by natural selection, even without genetic replication mechanisms, under specific conditions. Their results carefully distinguish selective forces such as conformity bias, operating on (the adoption of) individual cultural items, from natural selection, which operates on a population level.

---

[12] Tëmkin and Eldredge (2007) suggest that their reconstruction of instrument traditions is a demonstration of inaccuracy, not just of the reticulated shape of the tree of technology. However, they conclude the former from the latter: low phylogenetic signal in the data is taken as an indication that cladistic techniques misrepresent the true historical relations. This need not be the case, since the (many) homoplasies in a low-CI cladogram could be taken to represent horizontal transmission. The question is rather whether similarities known to be due to horizontal transmission are correctly reconstructed as homoplasies. The more consistent accuracy tests in (Eerkens et al. 2006) likewise focus on how simplified transmission systems lower consistency indexes to that obtained for a random data set. But again, inaccuracies might also arise for acceptable consistency indexes, because of specific individual-level mechanisms at work in the design and use of technology.

Their results do not automatically extend to *material* culture. Many advocates of theories of cultural evolution, including Boyd and Richerson, define culture so broadly that it includes technical artefacts as well as religious systems and hunting techniques. Yet artefacts might have some distinctive features, like the aforementioned strong modularity and material constraints. Moreover, many artefacts are embedded in multiple use practices: one person's hunting knife may be another person's razorblade. To show that the results of Boyd and Richerson extend to material culture, one would need to study systematically how their various selective forces may be affected by these features, and which supplementary forces might be needed. Only after these changes have been properly incorporated in their dual-inheritance models, it could be studied under which conditions, if at all, populations of artefacts evolve.

There have as yet been no attempts at this modelling by evolutionary archaeologists.[13] My point is not that these models cannot be constructed, or that their results will be negative. Rather, it is that without these models (or others like them), it cannot be said which mechanisms are responsible for the historical relations between artefact types which would be 'revealed' by a population-level phylogenetic reconstruction. Given the availability of software packages like PAUP, it is possible to generate such a reconstruction, which has every appearance of a result in archaeology. Yet explaining the shape of the cladogram is not be a gruelling job, as in the individual-level case, but – at least at present – a shot in the dark.

5. Conclusions

Constructing cladograms from morphological data seems a discipline-neutral operation. Applicability results in archaeology have indeed shown that parsimony-based methods (and the software in which it is implemented) travel unscathed from biology to the study of technical artefacts. I have argued, however, that the resulting reconstructions of phylogenetic traditions do not travel as well; a meaningful result in one disciplinary context may be of dubious value in another, although it has been

---

[13] There have been attempts to connect population-level mechanisms to (types of) artefact features. In particular, natural selection was claimed to operate on functional features and drift on stylistic features (Dunnell 1978). This function-style distinction is highly problematic (Hurt and Rakita 2001). Moreover, in the absence of models of relevant selective forces operating on functional and stylistic features, claims about mechanisms operating on populations of these features are premature.

gained by applying the same method. Disciplinary context affects phylogenetic analyses of tool traditions in several ways. First, they are supposed to contribute to explanation of the archaeological record rather than classification, since the latter is thought to be a matter of prior stipulation. Second, there is an unresolved, implicit ambiguity in the explanations sought – individual-level or population-level – which affects both the construction of cladograms and their interpretation. Third, on either explanatory level, cladograms contribute little to outstanding questions: on the individual level, few relevant mechanisms can be excluded from consideration on the basis of phylogenetic reconstructions, and the operation of some mechanisms may be systematically overlooked; on the population level, interpreting cladograms is mere speculation as long as it has not been established which mechanisms operate on classes of artefacts.

These conclusions are critical about existing work on tool traditions. It does, however, identify what I regard as some real research challenges and choices faced by evolutionary archaeology. Moreover, the conclusions do not entail that phylogenetic analysis cannot be applied successfully and productively outside of biology. Such applications may, for instance, be limited to extensions of the comparative method to studies of cultural evolution; there, reconstructions of tool traditions do not contribute directly to explanations, but they are relevant for checking dependencies in data sets and assessing the strength of correlations.

Another suggestion concerns the method for phylogenetic reconstruction. The current preference for parsimony-based methods, combined with the idea that cladograms can be constructed before they are interpreted, perhaps needs to be reconsidered. Alternatives are approaches based on Bayesianism, or on model-selection approaches to statistical inference. These methods are, to a large extent, ordinally equivalent to parsimony, i.e., they generate the same reconstructions (Sober 2008, Ch.4). *Generating* reconstructions, however, may be unproductive in archaeology, given the superabundance of mechanisms. Instead, one might try to *test* explanatory hypotheses by means of phylogenetic analyses. This reverses the two-step procedure proposed by evolutionary archaeologists and aligns with work in biology, in which the relevance of phylogenetic analysis for hypothesis testing was shown (e.g., Huelsenbeck and Rannala 1997; Johnson and Omland 2004). It would be worthwhile to examine whether this relevance can be extended to archaeology, along with the analytic methods used.

References

Basalla, G. (1988) *The Evolution of Technology*. Cambridge: Cambridge University
Press.

Bentley, P.J., ed. (1999) *Evolutionary Design by Computers*. San Francisco: Morgan
Kaufmann.

Boyd, R. and P.J. Richerson (1985) *Culture and the Evolutionary Process*. Chicago:
University of Chicago Press.

Collard, M., S.J. Shennan and J.J. Tehrani. (2006) "Branching versus Blending in
Macroscale Cultural Evolution", in: Lipo et al. (2006), pp.53-63.

Darwent and O'Brien (2006), "Using Cladistics to Construct Lineages of Projectile
Points from Northeastern Missouri", in: Lipo et al. (2006), pp.185-208.

Dunnell, R.C. (1978) "Style and Function", *American Antiquity* 43: 192-202.

Eerkens, J.W., R.L. Bettinger and R. McElreath (2006), "Cultural Transmission,
Phylogenetics and the Archaeological Record", in: Lipo et al. (2006), pp.169-
183.

Harvey, P. and M. Pagel (1991) *The Comparative Method in Evolutionary Biology*.
Oxford: Oxford University Press.

Huelsenbeck,.J.P. and B. Rannala (1997) "Phylogenetic Methods Come of Age",
Science 276: 227-232.

Hull, D. (1988) *Science as a Process*. University of Chicago Press.

Hurt, T.D. and G.F.M. Rakita, eds. (2001) *Style and Function*. Westport, CT:
Greenwood.

Johnson, J.B. and K.S. Omland (2004) "Model Selection in Ecology and Evolution",
Trends in Ecology and Evolution 19: 101-108.

Laland, K. and G. Brown (2002), *Sense and Nonsense*. Oxford: Oxford University
Press.

Laurence, S.L. and E. Margolis, eds. (2007) *Creations of the Mind*. Oxford: Oxford
University Press.

Lewens, T. (2004) *Organisms and Artifacts*. Cambridge, MA: The MIT Press.

Lipo, C.P., M.J. O'Brien, M. Collard and S.J. Shennan, eds. (2006) *Mapping Our
Ancestors*. Piscataway, NJ: Transaction.

Mace, R., C.J. Holden and S. Shennan, eds. (2005) *The Evolution of Cultural
Diversity*. London: UCL Press.

Mayr, E. (1981) "Biological Classification", *Science* 214: 510-516.

O'Brien, M.J. and R.L. Lyman (2000) *Applying Evolutionary Archaeology*. New York: Kluwer Academic/Plenum.

O'Brien, M.J., J. Darwent and R.L. Lyman (2001) "Cladistics Is Useful for Reconstructing Archaeological Phylogenies", *Journal of Archaeological Science* 28: 1115–1136.

O'Brien, M.J. and R.L. Lyman (2002) "Evolutionary Archaeology: Current Status and Future Prospects", *Evolutionary Anthropology* 11: 26–36

Renfrew, C. and P. Bahn (2004) *Archaeology*. 4th ed. London: Thames & Hudson.

Ridley, M. (2004) *Evolution*. 3rd ed. London: Blackwell.

Saviotti, P.P. (1996), *Technological Evolution, Variety and the Economy*. Cheltenham: Edward Elgar.

Sober, E. (1988) *Reconstructing the Past*. Cambridge, MA: The MIT Press.

Sober, E. (1991) "Models of Cultural Evolution", in: P. Griffits (ed.) *Trees of Life*. Dordrecht: Kluwer, pp.17-40.

Sober, E. (2008) *Evidence and Evolution*. Cambridge: Cambridge University Press.

Spencer, M. E.A. Davidson, A.C. Barbrook and C.J. Howe (2004), "Phylogenetics of Artificial Manuscripts", *Journal of Theoretical Biology* 227: 503-511.

Sterelny, K. (2006) "Memes Revisited", *British Journal for the Philosophy of Science* 57: 145-165.

Swofford, D.L. (2002) *PAUP: Phylogenetic Analysis Using Parsimony*. Sunderland, MA: Sinauer.

Tëmkin, I. and N. Eldredge (2007) "Phylogenetics and Material Cultural Evolution", *Current Anthropology* 48: 146-153.

Venditti, C. and M. Pagel (2008) "Speciation and Bursts of Evolution", *Evolution: Education and Outreach* 1: 274-280.

Ziman, J., ed. (2000) *Technological Change as an Evolutionary Process*. Cambridge: Cambridge University Press.