

# History Written in Stone: Evolutionary Analysis of Stone Tools in Archeology

Michael J. Shott

Published online: 11 June 2011  
© Springer Science+Business Media, LLC 2011

**Abstract** Populations of living things evolve over time, but do other things? Evolution involves transmission, be it of genes, ideas, or designs. What is transmitted, how and by whom, influences tempo and mode of evolution. In recent years, archeologists have applied evolutionary logic and processes to their study of things made and used by ancient people. Despite differences in subject units and in modes and patterns of transmission, evolutionary processes and the transmission modes that accompany them are worth seeking in archeological data. Stone spear points are abundant in the archeological record, yet we lack a theory to explain the creation, duration, and divergence of point types. Evolutionary studies of New World Late Pleistocene Paleoindian points are a step toward such theory, but limit the form of data and the evolutionary processes considered. An alternative in the study of Paleoindian points is geometric morphometric methods that do not constrain how point size and form are characterized nor assume branching divergence between taxa. Evolutionism should not dominate archeology, but it should become a major area of research within the field.

**Keywords** Points · Evolution · Geometric morphometrics

Archeology is a thing of cycles. Nineteenth century archeology cycled between ordering past time, exemplified in Thomsen's Three-Age system, function or adaption, and metaphysical evolutionism (Trigger 1996:121–138, 315–384). Culture history arose anew early in the twentieth century. Then mid-century processualists invoked function

to explain patterns of evidence. More recently still, culture history returned, clad in the armor of dual-inheritance theory and brandishing a cladistic lance. This incarnation denies much role to function yet severely constrains social process; it is culture history on mathematical steroids. For nearly two centuries, then, archeology has cycled between history, function, and evolution.

In each cycle, the dominant view largely precluded or denied alternatives; the archeological record was evidence either of passing style, function, or progressive evolution, rarely of all in complex interaction. As a result, archeology's intellectual history is a chronicle of cycling fads, not steady progress. We keep exploring only to arrive where we started, but never knowing the place any better. Unresolved cycling imparts a Manichaeian quality to archeological thought, a tendency to advocate views categorically opposed to others rather than to reconcile or integrate them.

It may be time for new approaches. Archeology should consider at once the roles of history, function, and evolution in the trajectories of change in past cultures and their components. Doing so requires new theory that Clarke (1968) adumbrated but that remains badly underdeveloped (Shott 2010a) and methods that are new, at least to archeology. This paper argues that archeology should consider an evolutionary approach to the study of artifact lineages that are common in the material record. That approach requires some new theory, whose broad outlines can only be sketched at this point, and methods of morphometric analysis suitable to the development and testing of that theory.

## History from Stone

In their study of the past, historians can choose among sometimes contradictory texts. But we have only one

---

M. J. Shott (✉)  
Department of Anthropology and Classical Studies,  
University of Akron,  
Akron, OH 44325, USA  
e-mail: shott@uakron.edu

archeological record, so we have no choice but to try to read it as best we can. All parts of the archeological record preserve information about past cultures, if we have the wit to interpret them. Some parts, however, recommend themselves to the task better than others. One is the chipped stone tools that archeologists often turgidly call “projectile points” and that lay people know as “arrowheads” (even if many of them tipped lances or spears, not arrows). “Point” is a suitable label, because these artifacts bear relatively sharp tips that were fashioned for the purpose of impaling and disabling prey, and were themselves hafted as the pointed tips of shaft weapons like spears and arrows.

Points had essential characteristics whose expression required considerable effort and elaboration; thus, a good deal of joint technological, morphological, and perhaps historical information resides in the size and form of points and the description of the types that they may comprise. Points were pointed (obviously!) for target penetration and were relatively thin to accommodate both flight and penetration. They were parts of larger assembled weapons, so were attached to shafts or foreshafts by varieties of mastic, lashing, and compression into slotted or socketed hafts. Different types of hafting imposed different requirements on the size and shape of points’ haft sections. Besides hafting and wounding by penetration, many points were designed with the secondary goal of durability by resistance to fracture. Opposing imperatives, for instance thinness for penetration and thickness for durability, made some points complex compromises between competing needs.

Hafting and performance have to do with how points served their users. They serve archeologists as well, because points were made of imperishable stone, so persist even though their shafts long since vanished. In the Americas, points were made in numbers practically beyond measure—even now, after centuries of rampant, uncontrolled collection (LaBelle 2003; Shott 2008a; Wilson [1899:233–239] documented collection on the scale of “wagonloads” [1899:234], more than a century ago), known points alone number in the millions from a record that surely numbered in the billions—and according to designs that changed over time, sometimes almost imperceptibly and sometimes rapidly. Whatever the abundance of points and the richness of their types elsewhere, enough were made in the relatively few millennia of New World prehistory to occupy us indefinitely. In global archeological perspective, the Americas are Point Central.

Imperishable and abundant, points obviously are popular subjects of study for the combined technological, morphological, and historical information they encode. Points fairly bristle with information if we can learn to read it in them. Among that information is the intrinsically historical character of the aggregate that they form over long periods

and the tendency for types to arise, flourish for intervals, then to disappear. Stratigraphic excavations and association with radiocarbon dates established the broad outlines of regional sequences of point types. For decades, therefore, archeologists used point types as convenient markers of gross time, typically intervals spanning millennia like “Early Archaic” or centuries like “Middle Woodland,” and sometimes as markers of cultural affinity. At the same time, points obviously represent important aspects of prehistoric behavior. Their presence in a deposit signifies, at a minimum, episodes of hunting or similar activities like armed conflict proportional in some way to their number. Thus, points jointly register time and activity, perhaps signifying “Early Archaic” hunting or “Middle Woodland” warfare.

Curiously, in a subject so abundant and diverse, however, archeology has devoted little effort to explaining the point sequences it constructs. We know that type A precedes type B precedes type C. But we cannot explain why type A persisted for as long or briefly as it did, why it fell from favor to be replaced by type B, or perhaps why it persisted in parts of its historical range while B arose in others or whether type A exhibited stasis and was replaced by unrelated type B or changed, evolved, morphed (depending upon your verb of choice) into type B (or into related types B<sub>1</sub> and B<sub>2</sub>). We know little about the relationship between the technological or morphological integrity of type A and its production complexity or range of materials from which specimens of it were made, i.e., its tendency toward variation and change. That is, archeology lacks a theory of the historical behavior of point types distributed across time, a sad state for so abundant and diverse a class and roughly comparable to a physics that lacks a theory of the atom.

### Do Points Evolve? If So, How?

Whatever the many causes of variation in points, one way to study and explain that variation is in evolutionary terms. The approach requires clearly defined units and heritable variation among them. Evolution is at once process and result, the thing that happens and the consequences of its happening. In this sense, the things that vary and upon which selection acts to produce evolution can be people, pea plants, pots, or stone tools. Objections to evolution applied to things would be valid if sexual reproduction were necessary to evolution, if evolutionary archeologists claimed animate properties for inanimate objects, or if selection pertained only to living things. Artifacts may not reproduce themselves but, like individuals within taxa and taxa within lineages, they may be products of inheritance subject to selection, and they (or at least sets of them) can be ordered in time (Apel and Darmark 2009; Shott 2008a).

In evolutionary studies, archeology borrows heavily in both theory and method from other fields, notably paleobiology. In organic evolution, traits are passed down, not up, across or diagonally; inheritance is vertical and taxa are branching. Cultural transmission can be reticulate, traits being transmitted horizontally and obliquely as well as vertically. An irony of the current moment is that paleobiologists (e.g., Tëmken and Eldredge 2007) seem more skeptical of the prospects for an evolutionary archeology than do archeologists (e.g., Bettinger 2009). Their skepticism owes in large measure to the specter of reticulate inheritance and horizontal transmission in things made, used and emulated by people. However, similar transmission mechanisms may be in the cultural and natural worlds, inheritance processes are distinct and vastly more complex in culture. The nature of cultural transmission is a serious complication but not an insurmountable barrier to the application of evolutionary theory; in fact, we probably should assume until proven otherwise that cultural transmission is reticulate. Archeologists and other anthropologists who apply evolutionary theory to material things have developed methods to detect and measure, and thereby control for, horizontal transmission (Currie et al. 2010) or at least to identify the conditions that make historical inferences robust to its effects (Nunn et al. 2010).

Whatever the validity of evolutionary theory applied to material culture, most applications involve expressive material culture like basketry that preserves too poorly in the archeological record to contribute significantly to our knowledge of the past. For instance, Borgerhoff Mulder et al. applauded “evolutionary archeology” (2006:55) studies that mostly involved materials exceedingly rare in the record! Among imperishable materials, evolutionary applications seem more common in pottery than in stone, and almost always in decorative, not functional, elements of pottery. (Whatever the merits and abundance of ceramics, fragmented assemblages of sherds are poorly suited to the morphometric approach to historical study sketched below.) Ironically, the least perishable, most abundant and most ancient craft—stoneworking—until recently was poorly represented in evolutionary studies (Shott 2008a). That situation is changing for the better, but my purpose is less to review evolutionary analysis in stone tools than to explore some of its prospects and challenges.

### What Kind of Units Are Points and Types?

Before the paleobiological revolution, biologists knew that individuals were the sole unit of selection (e.g., Williams 1966). Ever since, biologists have disputed the size and nature of selection units and the evolutionary processes that selection produced (Gould 1980). Whether genes are

individuals, and despite continuing debate about whether taxa are individuals, no biologist would argue that teeth or dorsal fins were individuals. Every biologist knows also that individuals, however defined at whatever scale, vary but only larger populations evolve. Thus, there is dispute within biology about what processes act at what level on which units, but comparatively little ambiguity about unit definition.

There is greater ambiguity in archeology. People undeniably are individuals and legitimately defined types may be individuals, but are points individuals? Points were designed and made as parts of larger apparatuses; no one ever threw an unhafted point at a target any more than soldiers today throw unchambered bullets. In this sense, points cannot be individuals. But points were made as individual objects and their size and form were subject to makers’ selection before hafting to shafts. The shafts and other elements of the larger apparatuses in which points served usually were perishable and are not preserved; practically, then, points are individual objects. Ambiguity resides in the tension between the strict view that points were parts, not wholes, and the practical recognition that they may be treated as individuals.

If, like O’Brien and Lyman (2003:36), we view artifacts as part of people’s phenotypes, a further ambiguity arises. Eye color, ability to roll one’s tongue, and the like are parts of the phenotype that mostly are constant and unchanging individual attributes. Once a blue-eyed tongue-roller, always one. But one individual could, often probably did, make and use hundreds of points in (usually) his lifetime. Collectively, all may be part of the individual’s phenotype, but surely there was variation among them no matter how constant the individual’s blue eyes and tongue rolling. With respect to points, the individual’s phenotype must be characterized by range and pattern of variation among the points, some of which owes to chance or drift, time, experimentation, and adaptation, raw material and the like. In this sense, the points that comprise an individual’s phenotype flow through time in a dense, near-continuous stream. Individuals did not live and die in one place but instead ranged across landscapes. Therefore, each individual’s phenotypic point population is distributed in many assemblages, even though each individual point occurs, trivially, only in one. The flow of points as phenotype is scattered in many channels.

The similar properties of similar points produced by contemporaries and those earlier and later mean that, in all but the rarest circumstances, it is similar folly to attempt to distinguish points made by different individuals. In most archeological contexts, it is equal folly to attempt chronological resolution even so finely as to individual lifetimes; most archeological deposits are time-averaged over coarser scales (Shott 2008b; 2010a, b; Stern 1993). This condition

makes typical point assemblages free combinations of specimens made by many people over considerable time; they are the compound of synchronic variation between individuals, of variation over individuals' lifetimes, and of variation over many lifetimes—each one's flow distributed in many channels.

The vagaries of deposition and sampling introduce further complications. Not all point-bearing deposits that formed are preserved, but instead redeposited, nor are preserved in accessible contexts, but instead deeply buried. Limited sampling breaks the flow of points through time into discrete units that vary in number and duration. Some units can be ordered in time stratigraphically or by absolute date, but some cannot. As a result, the flow of points through time may form unbroken historical lineages of evolution when perfectly sampled. But in the material record, sampling is far from perfect. Flows pool into units qua assemblages, which themselves are the composites described above, not nearly perfectly time ordered and individually separated flows. Assemblages themselves sample the flow of time quite imperfectly, arbitrarily breaking it into slices separated by gaps, both of which vary widely in duration. Point assemblages, therefore, are time-averaged compounds that combine the partial flows or pools of many individuals distributed across much time.

Yet fossil assemblages also are time-averaged. The individuals that comprise them, often separated by many generations, may exhibit diachronic variation freely combined with synchronic or polymorphic variation in the taxa defined from the assemblages (Newell 1956:68–69). The fossil record is as imperfectly sampled in time and space as is the archeological record. In both archeology and paleobiology, then, individuals occur in time-averaged assemblages. Types or taxa are defined from the aggregate properties of assemblages. In both fields, the flow of time and specimens through it is cut into frozen slices of varying length and breadth. Each slice is treated as a unit, not a section of a flowing stream. It is from such frozen slices that our view of evolution's flow must arise.

In this perspective, points bear especially salient resemblance to invertebrate fossils, which played an important role in the rise of modern paleobiology (Sepkoski 2009:15–16). Like invertebrate fossils, in the Americas at least bifacial points are superabundant. This quality minimizes sample-size problems and makes points highly suitable for quantitative analyses not always valid or practical in other artifact classes (e.g., ceramics, the vast majority of which occur as fragmented sherds, not integral vessels). Points also are “single rigid structures” (Polly 2004:1; if, admittedly, originally parts of larger complex devices) amenable to geometric modeling as Polly described.

Historical sequences require chronological control. Time order is given in stratigraphic sequences and by points

associated with absolute dates. There are many excavated sequences but the numbers of points from those sequences are rather modest (perhaps unlike invertebrate fossils), nor is their co-occurrence in stratified sequences evidence of historical continuity in a set of points. Association with absolute dates also involves modest numbers of points. Ideally, the best sequence data come from direct dating of many points. Where obsidian points are common, hydration dating, warts and all, is one source of such data. Otherwise, thermoluminescence dating of points made of heated chert or other stone is the best candidate (e.g., Wilhelmsen 2001).

Until finely resolved sequences that include hundreds or thousands of points are constructed—a task that will take some time—archeologists must emulate paleobiologists. We must first define point types (either “intensionally” [empirically] or “extensionally” sensu O'Brien and Lyman 2003:138) from imperfect data that encompass the considerable synchronic and diachronic variation noted above. Then from what chronological data exist, we must order the types in time. Then we must document historical lineages in the sequences. Then focusing on those lineages, we must describe in detail the rate, pattern, and magnitude of typological change along the sequences. Then we must try to explain the sequence of change.

### If Points Evolve, How Do We Measure the Process?

If points evolve, they do so in response to joint functional and social factors. The functional imperatives of points are known in the broadest terms: flight characteristics, thrust or kinetic energy, penetration, and perhaps durability (Chesier and Kelly 2006). Hughes (1997:349–350) identified accuracy, range, and penetration as important performance criteria; Wilhelmsen (2001:112–115) added aerodynamic properties related to flight stability. Variation in performance requirements can be measured, among other ways, by characteristics like tip acuity, bilateral plan symmetry, and cross-section form and perimeter (Hughes 1997; Wilhelmsen 2001; see also Beck 1998:25; Chesier and Kelly 2006; Churchill 1993; Collins 2007:72; Ratto 2003: Table 5.2a,b; Shott 1993). From hafted ethnographic and archeological specimens of known status, Hughes distinguished points of thrusting spears, darts, and arrows, which occupied somewhat overlapping segments of continua in mass, cross-section area and perimeter (1998:Tables III–IV; see also Ratto 2003:215–216).

Hollywood movies to the contrary, it is unlikely that points signaled ethnic identities like the “Apache” arrowheads that John Wayne, teeth gritted, pulled from his own shoulder. Whatever signaling of ethnic status probably resided in larger, more visible segments of projectiles like their shafts (Sinopoli 1991). Yet points may have been

subject to Boyd and Richerson's socially determined transmission processes, although the evidence sometimes is equivocal (e.g., Shott 1997). As little as we know of points' functional performance, we know even less of their social one. One future challenge is to identify traits in points of social significance.

### If Points Evolve, How Do We Analyze the Process?

Points have been used to test transmission theory, on the logic that particular patterns of attribute covariation and rates of change in attributes are most consistent with one or another of Boyd and Richerson's (1985) models. An earlier essay reviewed important studies (Shott 2008a:150–151), unnecessary to repeat except in selected respects. Bettinger and Eerkens (1999) interpreted different patterns of attribute correlation in Rosegate points, probably arrowheads, from the western and central Great Basin as products of distinct transmission modes. Base width and weight correlated in the central but not the western region. The complex inference was that people in the central region adopted bow-and-arrow technology as an integrated whole, a package deal (therefore, indirect bias), whereas people to the west assimilated the technology piecemeal (therefore, guided variation or trial-and-error adaptation).

Such studies suggest that the application of transmission theory to stone tools may be productive. Yet from their agent-based model, Mesoudi and O'Brien argued that the same attribute patterns were consistent with other transmission models besides indirect bias, i.e., an equifinality thesis. They concluded that "it will be difficult in practice to distinguish between different forms of cultural transmission in the archeological record *when relying solely on attribute correlations or measures of variation*" (2008:640, original emphasis). Polly (2004) reported similar difficulties in distinguishing among transmission models or selective regimes in data of the kind and quantity common in the archeological record; Borgerhoff Mulder et al. (2006) implicitly situated such transmission models at the micro-evolutionary scale and focused instead on distinguishing vertical and horizontal transmission patterns at the macro-evolutionary scale. Mesoudi and O'Brien's study also linked synchronic and diachronic attribute correlations with adaptive constraints, using Sewall Wright's "adaptive landscape" concept. If Mesoudi and O'Brien are correct, then traditional analysis is insufficient and simulation of transmission by different modes under different adaptive constraints is equally important.

Easier said than done. The "adaptive landscape" of so humble a class of so small objects as points may be remarkably complex. (If it is not, but instead unimodal, then Mesoudi and O'Brien [2008:641] argue that guided

variation always is the best solution.) If, therefore, we believe or at least suspect that some variation in point size, form, and time-space distribution may owe to transmission process, then we assume complex adaptive landscapes de facto. There is, first, the reconciliation of competing performance requirements of flight, penetration, and durability (e.g., Chesier and Kelly 2006; Hughes 1997). Second is the relationship between performance requirements and prey characteristics like density, effective range, flight behavior, body size, etc. Third are social constraints imposed by, for instance, labor specialization or even social norms that oppose specialization (e.g., Cundy 1989). There may be other factors, which we will know only when we think sufficiently long and hard about the problem. Points emerge as remarkably complex objects whose reading poses stiff challenges.

When possible, knowing the characteristics of different transmission modes in historical point sequences is a valuable development. At the same time, it is essential that inference to transmission mode does not become an end in itself. Returning to Bettinger and Eerkens (1999), it is good to know that guided variation characterized the western Great Basin mid-Holocene point sequence but indirect bias characterized the eastern sequence. But this answer merely begs another question: why should either transmission mode prevail where it did? Different transmission modes themselves can be favored in different circumstances of, for instance, population size or density, technological constraints or innovations, and sociopolitical contexts. As a result, inference to transmission mode alone is partial inference at best. Besides transmission theory itself, archeology requires a higher-order body of theory that specifies the conditions in which various transmission modes are apt to prevail.

Chronological resolution usually is coarse in the time-averaged deposits that are most archeological assemblages. Boyd and Richerson's transmission models involve social processes and contexts that can change at rates much faster than the span of most time-averaged assemblages. Given the coarse time resolution of assemblage data in which most points occur, they easily could accommodate combinations of transmission models. In time-averaged archeological data, then, detailed transmission models may be difficult to distinguish not only because of the equifinality noted above but also because they act over finer time scales than do the data to which they apply.

### History from Points: The Paleoindian Case

Evolutionary studies of stone tools are on the rise. Most involve Old World tools (e.g., Archer and Braun 2010; Castiñeira et al. 2007; Clarkson et al. 2006; Grosman et al.

2008; Lycett 2009; Sumner and Riddle 2008) that probably did not function as points (but see Apel and Darmark 2009). Collectively, these studies attest to the promise of lithic evolutionary analysis. Here, I confine treatment to current and new approaches in New World Pleistocene archeology.

In developing archeological theory of points, from a New World perspective, it is fitting to begin at the beginning. The Paleoindian point sequence starts with Clovis and diversifies into Goshen and Folsom on the Plains, to Gainey and Barnes/Cumberland in the Midwest, perhaps directly to Dalton in the mid-South, to Suwannee, Simpson and other unfluted lanceolates in the deep South, and to Colas de Pescado in Latin America. Sequences are well described and, as descriptive labels, types are convenient markers for complex patterns of variation. But we must attend to the properties of these sequences as streams, not ice blocks. We do not know if Clovis morphed by continuous degree into later types (phyletic gradualism), if types branched off fully formed while Clovis persisted (cladogenesis), or if later types are unrelated replacements of Clovis. In this state of ignorance, archeologists have begun to test methods to distinguish among the possibilities.

Functional and historical dimensions of variation must be considered in any point study. In the Paleoindian case, functional dimensions include not only flight characteristics, penetration, and durability but also launching device. Collins (2007:76–79) summarized data on use–wear type, placement and pattern, and equivocal direct evidence for atlatls to conclude that Clovis points probably were hand-held lances, although he did allow the possibility that they were darts launched by atlatls. Their similarity in size and form recommends Collins’s judgment to Gainey/Bull Brook points of eastern North America. (The status of Suwannees, Barnes/Cumberlands and Colas de Pescado is less certain.) In contrast, most archeologists believe that Folsom and probably Dalton points were launched by atlatls. If Collins is correct, then later fluted points like Folsoms and later types may be skeuomorphs adapted from a preexisting design to accommodate a new launch device. If so, then future studies, either cladistic as summarized below or geometric morphometric as advocated later, must take into account variation that corresponds to launching device.

#### O’Brien and Lyman

O’Brien and Lyman (2003) was the first major study of evolutionary paths in points, involving a large sample of early Holocene types from the American southeast. In context, the study was stimulating but inspired certain reservations. O’Brien and Lyman modeled three-

dimensional points as two-dimensional objects, using eight mostly nominal and ordinal variables (2003:152). Despite indifference to “whether characters are quantitative or qualitative” (2003:144), their commitment to paradigmatic classification prompted O’Brien and Lyman to reduce naturally continuous data to interval and lower scales. Unfortunately, at least three of their variables conflate original design with the negative allometry of resharpening, i.e., code specimens for degree and pattern of resharpening experienced (Shott 2008a:148–149).

Types (“taxa”), not individual points or empirically assemblages of them, were O’Brien and Lyman’s unit of study. They were right to question (but not necessarily reject) the empirical basis of type definitions. Problems include lack of precision and consistency in description, disagreement about what constitutes essential characteristics, and change in type definition as more empirical samples are found. Surely classification can be improved. But their paradigmatic alternative—the a priori definition of classes into which empirical specimens are placed according to their sets of characteristics—itself amounts to a reproduction of empirical types. O’Brien and Lyman argue that character lists and paradigmatic classes defined from them “are not empirical...they can only be created” (2003:139). But characteristics are selected from infinite possibilities and types defined that resemble empirical specimens only after examination of empirical samples. The paradigmatic approach does not avoid the problem of empirical type definition. In their study, moreover, although variables were chosen (“based on expectations as to which parts of a projectile point would change the most over time as a result of transmission” [2003:150]), the variables seemed selected to describe general qualities of plan form, haft element, and fluting; the classification was not obviously informed by the detailed examination of point performance or historical criteria noted above. That is, it was extensional as much as intensional.

Whatever the virtues of paradigmatic classification, in the 621 points of O’Brien and Lyman’s study, it produced 491 types *qua taxa*, an average of 1.26 points per type (2003:157). Such types are nearly as unique as snowflakes. Then they confined analysis to the 83 points that formed 17 larger types. In the process, a group of more than 600 specimens was reduced to an analytical set of fewer than 100, which seems inefficient in the use of empirical data if not unrepresentative. Whatever the imperfections of traditional typology, O’Brien and Lyman’s “intentional” types freely combined form, technology, fluting, size, and degree of resharpening and freely cross-cut traditional types (2003:156). This failure to replicate “extensional” types does not automatically discredit either approach, but does suggest that we should be more careful in how types are defined. Among other things, resharpening effects can be

minimized if not entirely removed by the simple expedient of confining characterization and analysis to haft elements, which are far less susceptible to resharpening than are blades and whole-object size and form.

Finally, cladistics found phylogenies for the unordered characters of the reduced point-type sample. Cladistics assumes branching divergence and usually is constrained to change only one state per step in ordinal data. These assumptions may or may not be faithful to the nature of transmission and selection in points. Assuming validity of the phylogeny (2003:169), O'Brien and Lyman make it the conclusion of their analysis. As description, this is fair enough, but any historical inference reached or evolutionary trend suggested, certainly one so detailed and complex as theirs, begs for broader contextualization and explanation.

### Other Cladistic Approaches

Buchanan et al. used two-dimensional images of a continental sample of fluted points. They defined landmarks and used inter-landmark distances as dimensional measures. They used two different methods to attempt to control for resharpening allometry (Buchanan and Collard 2007:372–373; Buchanan and Hamilton 2009:284), both critiqued elsewhere along with interpretations of results (Shott 2009). Assemblage was the unit of analysis, and mean “size-free” dimensions by assemblage were reduced to ordinal or interval-scale variables, treated as ordered states for cladistic analysis. The dataset included caches, and what most archeologists would call “kill” (e.g., Naco, Lehner) and “habitation” assemblages that varied considerably in size and probably accumulation span, which are important sources of assemblage variation (Shott 2010b).

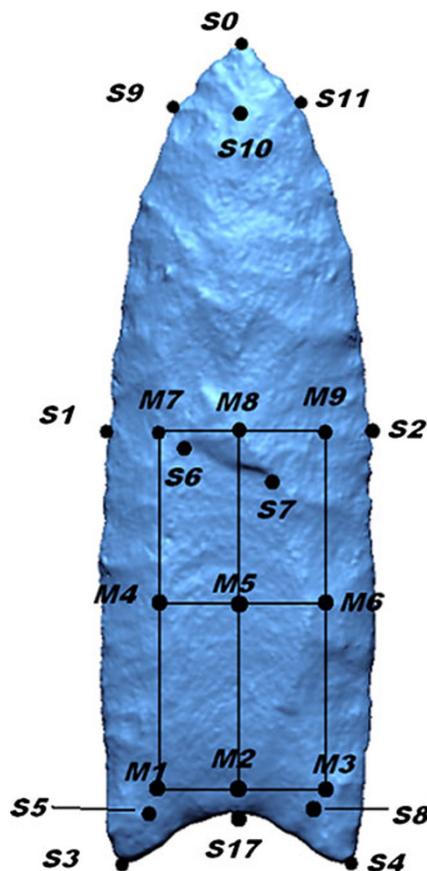
Here, the most salient points are that assemblages were units of analysis and that Buchanan et al. sought historical information in point data. Assemblages can be legitimate units of analysis, but free combination of caches, kills, and habitations is questionable. Caches were deposited in virtual instants of time and might comprise products of one knapper. Kill assemblages probably also accumulated in very short intervals, even if points were contributed by several users, unless they occur at places suitable to repeated ambushes or drives. But what archeologists typically call “habitation” sites, assemblages accumulated over much longer and more variable intervals. Combining caches with habitations freely mixes instantaneous deliberate deposits with accumulations over intervals orders of magnitude longer. Assemblages can serve as units of analysis, but not all assemblages are equal in size or accumulation span. Considering the great variation implicated, assemblages should comprise units of analysis only when at least roughly comparable in context, span, and size.

Among other limitations, cladistic practice in archeology emphasizes discrete traits to the virtual exclusion of continuous ones, which alters the character of data and might influence results. Our discrete-trait focus prevails across the range of analytical scales from characteristics of individual artifacts (e.g., Buchanan and Collard 2007; O'Brien and Lyman 2003) to entire cultures (e.g., Bettinger 2009; Chatters and Prentiss 2005). Yet some paleobiologists argue that continuous traits are perfectly amenable to cladistic treatment (e.g., MacLeod 2002). Of course we should treat as discrete those traits that legitimately are, but we should not reduce the intrinsic properties of our often continuous data merely to fit convenient analytical practices developed in other fields; instead, we should consider inherently continuous traits equally with discrete ones, and then use existing or devise new methods for their joint study.

### Geometric Morphometrics

Whatever the advantages of cladistic analysis (cf. Shott 2008a; 2009), geometric morphometrics (GM) also is on the rise, mostly in resharpening studies (e.g., Shott and Trail 2010). GM is a set of methods that produce and analyze data in ways historical among others. Like cladistics, GM reduces the complex whole objects that are points to smaller sets of dimensions. Whether in two- or three-dimensional models, however, GM does not reduce naturally continuous data to lower measurement scales. Unlike traditional manual measurement of orthogonal dimensions, GM data are free of geometric constraints. When manipulated in CAD or similar software, resulting models allow easy and accurate measurement of variables otherwise difficult to measure (e.g., volume, surface area, centroid size, longitudinal and transverse section area and perimeter [and variation in both at regular intervals along the axis], edge perimeter). Landmarks and other GM data can be placed to capture the functional traits that theory identifies (e.g., Hughes 1997; Wilhelmsen 2001) and the and technological traits that relate particularly to fluted points (Shott and Trail 2010; Fig. 1). Unlike cladistics, it is not an inherently historical approach that assumes mode, rate, or direction of evolutionary change.

MacLeod (2002:129–134) demonstrated the potential of GM in paleobiology. He defined what amount to modules—sets of landmark points that covary closely as quasi-discrete components (e.g., Klingenberg 2008) of larger wholes—in hypothetical phylogenies. Landmark  $x$ - $y$ - $z$  coordinates and inter-landmark distances are continuous values easily analyzed in GM routines like the program Morphologika (O'Higgins et al. 2009). MacLeod defined clusters of landmark configurations that both captured discrete characters in a time-ordered set of taxa and that reconstructed its



**Fig. 1** Landmark configuration for functional and technological analysis of North American Paleoindian points. Source: Shott and Trail (2010: Fig. 4)

evolutionary history (itself known in these hypothetical data, which thus serve as a control sample). Finally, he reduced continuous data to discrete characters and conducted conventional cladistic analysis that agreed closely with phylogenetic inferences reached in continuous data. Thus, MacLeod suggested that both continuous data and morphometric methods can reveal historical relationships in fossil lineages.

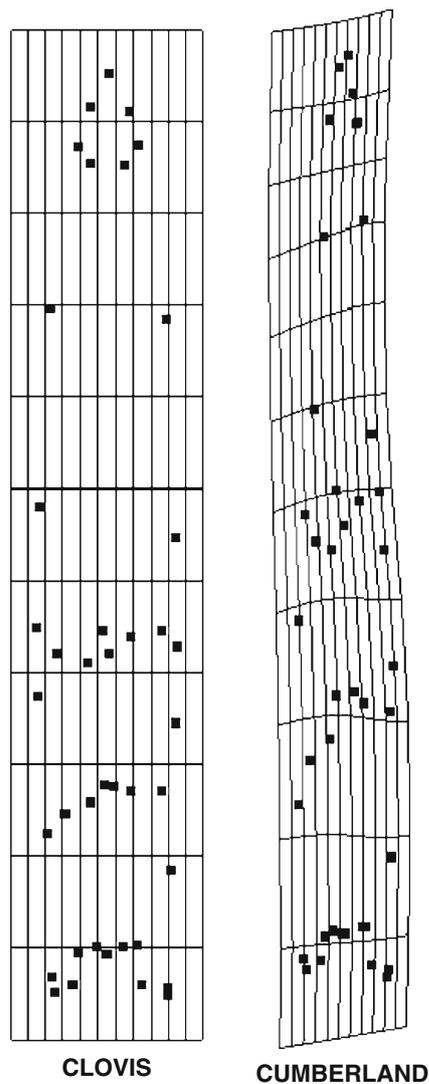
GM study of Paleoindian points might use types and assemblages differently as units of analysis. For many purposes, types should be units of analysis (Apel and Darmark 2009:16). That does not mean that types *as defined subjectively* or points casually assigned to types should be taken at face value. On the contrary, GM data can be used to define valid types, especially in haft modules rather than whole points to control for resharpening allometry. This can be done using standard cluster analysis, relative-warp cross plots (MacLeod 2002:131), canonical variate analysis, or other methods.

O'Brien and Lyman (2003:135) argued that empirical-type definition obscured historical relationships between Clovis and Dalton types, and that intensional paradigmatic classification and cladistic analysis alone could reveal it.

Yet GM characterization and analysis can reveal such relationships and is more faithful to the total form and continuous dimensions of points (i.e., does not reduce them to a small set of abstract formal properties and measures continuous variation in continuous terms). O'Brien and Lyman also argued that paradigmatic classification can define morphospace, "the multidimensional space encompassing the range of morphological variation of...taxa" (2003:140). So too can GM and landmark data; indeed, Gould (1991:420) considered cladistics and the paradigmatic classification consistent with it poorly suited to defining morphospace. Leaving aside the questions about O'Brien and Lyman's analysis noted above, GM, besides its unique virtues, can accomplish all that their approach claims.

For some purposes, however, cache assemblages may be valid analytical units for the properties noted above, either by defining points in a cache as a type or using context to test the validity of types defined otherwise. Among their virtues is that cache specimens rarely are resharpened, so resharpening allometry is mooted and entire points, not just haft modules, can be studied. Also, specimens in a cache often are of the same material, controlling this source of variation. Although it always remains an inference, cache points also may have been made by one person at essentially one time, controlling other sources of variation. Paleoindian point caches are difficult to find and therefore rare, yet a surprising number exist (Kilby 2008). It is worth at least exploring GM analysis in cache points.

Whatever the unit of analysis, and assuming some a priori information about the chronological distribution of types, historical or evolutionary analysis can take several routes. For instance, thin-plate splines and relative-warp plots are graphic depictions of pattern and scale of difference in landmark points between putative ancestral and descendant types (e.g., Shott and Trail 2010). Granted, they are graphics, not analysis, but scaled images of the deformation required to transform type A into type B can suggest hypotheses for rate, mode, and cause of change (Fig. 2). Some pictures are worth at least a few words. Software like MorphoJ (Klingenberg 2011) and Mesquite (Maddison et al. 2009) can test different transmission modes or other evolutionary processes in GM data, although they require a priori phylogenies, and there are ways to fit GM data to phylogenies that treat shape as a multivariate character and tests for historical signals in those data (Klingenberg and Gidaszewski 2010). It also should be possible to fit empirical GM trends to different models of adaptive landscapes (Bettinger 2009; Polly 2004). Moreover, GM data and perhaps inter-landmark distances allow points to be subdivided into modules or even smaller sets of data and dimensions. Resolving point types this way, we may find that evolution occurs much



**Fig. 2** Thin-plate splines (tps) of Clovis and Cumberland points, showing degree and pattern of deformation accumulated between types. Source: Shott and Trail (2010: Fig. 10)

differently or more rapidly in some segments than others, extending to multivariate GM data the approach that Morrow and Morrow (1999) demonstrated in the transition from Clovis to Colas de Pescado.

Whether microscale transmission processes can be identified in often-coarsely resolved archeological data, GM point data also can be used to explore macro-scale transmission. Again, the best examples are in paleobiology. Polly (2004) used computer simulation to model the evolution of complex morphology over long time spans in morphometric data on shrew tooth crowns. To constrain the pattern and magnitude of morphological change to the biologically realistic (i.e., to prevent morphometric landmark configurations to evolve or drift via stochastic simulation into unrealistic form [Polly 2004:3]), he calibrated covariance matrices to empirical data. (Channeling

Gould, Bettinger [2009:279] called “*Baupläne*” similar functional constraints at presumably higher levels than points, thus acknowledging constrained variation in cultural phenomena.) Polly then simulated evolution in simple (i.e., unimodal, one peak) and complex (i.e., multimodal randomly distributed peaks that varied in slope and height) adaptive landscapes under distinct selection modes: random fluctuation, directed or sustained trend, stabilizing, and drift. Morphologies evolved over long time patterned differently depending upon selection mode. Under random fluctuation, for instance, evolved morphologies were narrowly and randomly distributed near the origin (representing the starting morphology) of the plot of the first two axes of principal component space and divergence from starting morphology rose modestly over time while the range of morphologies rose monotonically (Polly 2004: Fig. 5). In contrast, directional selection yielded tightly clustered evolved morphologies distant from the origin of principal component plots with limited divergence from the selection trend and monotonic rise in divergence from starting morphology over time (Polly 2004:Fig. 7). Thus, Polly could distinguish the effects of different selection regimes and adaptive landscapes. He concluded (2004:22) that evolutionary trends and historical relationships between taxa can be identified in morphometric data under any mode except strong stabilizing selection.

It may seem a long stride from shrew teeth to points, but both are rigid solids amenable to morphometric characterization and whose landmark configurations are constrained by function, size, allometry, and other factors. Just as shrew teeth cannot grow too large or sprout needle-sharp peaks, points and their modular components cannot grow too large or small, cannot sprout multiple points on their edges or faces, cannot grow too thick or thin, and otherwise are functionally constrained (e.g., Hughes 1997; Wilhelmssen 2001.)

### Why do Points Evolve?

In a short essay crowded with many questions, “Why do points evolve?” is the \$64,000 winner. Like all important questions, this one is much easier to ask than to answer, and I can offer no remotely satisfactory one at this juncture. After all, an answer requires a body of theory that at once gives meaning to data and explains their pattern and variation by reference to external factors or internal properties. Physicists possess such theory to explain, for instance, the structure and behavior of atoms. Archeology, unfortunately, lacks theory to explain the historical properties—time–space range, tempo and mode of change, pattern, scale and direction of morphometric change itself wrought by complex interaction between

constraints, inertia and selective factors—in historical lineages. Lacking that theory, it is fatuous at present to attempt to answer the question posed.

Yet at least we might begin thinking about the problem. A host of subsidiary questions must first be posed and answered before sufficient theory emerges. Does time–space range of specific types or longer historical lineages vary with material, human population, environmental conditions, or change? Do tempo or mode of change vary with abundance of items, an equivalent of population size? Do rate and type of branching divergence depend upon the same factors, or perhaps others like raw material, complexity or length of the production process or artifact use-life? What external factors like population, sociopolitical organization, and complexity, constrain morphometric trajectories of change? What internal factors like material (again) or production-complexity (again) similarly constrain morphometric trajectories? Until we address such questions, we never will find the answers nor build the theory that can explain why and how points evolve.

If we cannot yet answer the \$64,000 or even the 99¢ questions, we can take at least a fool’s comfort in the possibility that methods like GM might reorient our thinking in ways that invite contemplation of the answers. In this way, needed theory and available method can forge a recursive relationship, advances in one stimulating reaction and revision in the other.

## Conclusion

However important systematic historical studies will be in archeology’s future—and I believe strongly that they should be—they must not preclude other approaches. We should continue to carry out micro-scale or short-term studies of adaptation or function in particular time-space settings, meso-scale studies of culture process, purely historical studies that focus on sequence and dating for their own sake, broad syntheses, methodological analyses, among others. But we should add to this customary mix studies of which the approach advocated here forms only a small part: the properties of macroscale or long-term historical lineages of material at all levels from trait, artifact, assemblage, to culture, following Clarke’s (1968) framework and recently revived by Bettinger (2009) and Chatters and Prentiss (2005) among others.

This is no brief for a reorientation of archeological research exclusively toward systematic historical studies; doing so would merely perpetuate the unproductive cycling between intellectual fads that characterized so much of archeology’s history and hampered its development. It is a truly modest proposal that archeology attend to intermediate and macroscale historical processes and develop the theory

and method to not merely describe and document, but to explain them. In view of their abundance, stone tools hold particular promise for this partial redirection of archeological thought and practice and transmission processes form one small part of a much larger theoretical apparatus. At this remove, we can only discern the faint outlines of that approach. Bringing it into sharp focus is a major challenge to archeology’s future.

## References

- Apel J, Darmark K. Evolution and material culture. *Curr Swed Archaeol.* 2009;17:9–26.
- Archer W, Braun DR. Variability in bifacial technology at Elandsfontein, Western Cape, South Africa: a geometric morphometric approach. *J Archaeol Sci.* 2010;37:201–9.
- Beck C. Projectile point types as valid chronological units. In: Ramenofsky A, Steffen A, editors. *Unit issues in archaeology: measuring time, space, and material.* Salt Lake City: University of Utah; 1998. p. 21–40.
- Bettinger RL. Macroevolutionary theory and archaeology: is there a big picture? In: Prentiss A, Kuijt I, Chatters J, editors. *Macroevolution in human prehistory.* New York: Springer; 2009. p. 275–95.
- Bettinger RL, Eerkens J. Point typologies, cultural transmission, and the spread of bow-and-arrow technology in the prehistoric great basin. *Am Antiquity.* 1999;64:231–42.
- Borgerhoff Mulder M, Nunn CL, Towner MC. Cultural macroevolution and the transmission of traits. *Evol Anthropol.* 2006;15:52–64.
- Boyd R, Richerson RJ. *Culture and the evolutionary process.* Chicago: University of Chicago; 1985.
- Buchanan B, Collard M. Investigating the peopling of North America through cladistic analysis of early paleoindian projectile points. *J Anthropol Archaeol.* 2007;26:366–93.
- Buchanan B, Hamilton MJ. A formal test of the origins of variation in North American early paleoindian projectile points. *Am Antiquity.* 2009;74:279–98.
- Castiñeira C, Cardillo M, Charlin J, Fernicola JC, Baeza J. Análisis Morfométrico de los Cabezales Líticos ‘Cola de Pescado’ del Uruguay. In: *Arqueometría, Segundo Congreso Argentino.* Comisión Nacional de Energía Atómica, Buenos Aires; 2007. pp. 1–7.
- Chatters JC, Prentiss WC. A Darwinian macro-evolutionary perspective on the development of hunter-gatherer systems in Northwestern North America. *World Archaeol.* 2005;37:46–65.
- Chesier J, Kelly R. Projectile point shape and durability: the effect of thickness: length. *Am Antiquity.* 2006;71:353–63.
- Churchill SE. Weapon technology, prey size selection and hunting methods in modern hunter-gatherers: implications for hunting in the Palaeolithic and Mesolithic. In: Peterkin G, Bricker H, Mellars P, editors. *Hunting and animal exploitation in the later paleolithic and mesolithic of Eurasia.* *Archeological Papers of the American Anthropological Association* 4:11–24; 1993.
- Clarke D. *Analytical archaeology.* London: Methuen; 1968.
- Clarkson C, Viniciu L, Lahr MM. Quantifying flake scar patterning on cores using 3D recording techniques. *J Archaeol Sci.* 2006;33:132–42.
- Collins MB. Discerning clovis subsistence from stone artifacts and site distributions on the Southern Plains Periphery. In: Walker R, Driskell B, editors. *Foragers of the Terminal Pleistocene in North America.* 2007. pp. 59–87.

- Cundy BJ. Formal variation in Australian spear and spearthrower technology. *British Archaeological Reports IS 546*. Oxford: Oxbow; 1989.
- Currie TE, Greenhill SJ, Mace R. Is horizontal transmission really a problem for phylogenetic comparative methods? A simulation study using continuous cultural traits. *Philos Trans R Soc B*. 2010;365:3903–12.
- Gould SJ. The promise of paleobiology as a nomothetic evolutionary discipline. *Paleobiology*. 1980;6:96–118.
- Gould SJ. The disparity of the burgess shale arthropod fauna and the limits of cladistic analysis: why we must strive to quantify morphospace. *Paleobiology*. 1991;17:411–23.
- Grosman L, Smikt O, Smilansky U. On the application of 3-D scanning technology for the documentation and typology of lithic artifacts. *J Archaeol Sci*. 2008;35:3101–10.
- Hughes SS. Getting to the point: evolutionary change in prehistoric weaponry. *J Archaeol Meth Theory*. 1997;5:345–408.
- Kilby JD. An investigation of Clovis caches: content, function, and technological organization. Unpublished Ph.D. dissertation, Department of Anthropology, University of New Mexico. Albuquerque; 2008.
- Klingenberg CP. Morphological integration and developmental modularity. *Annu Rev Ecol Syst*. 2008;39:115–32.
- Klingenberg CP. MorphoJ: an integrated software package for geometric morphometrics. *Mol Ecol Resour*. 2011;11:353–7.
- Klingenberg CP, Gidaszewski NA. Testing and quantifying phylogenetic signals and homoplasy in morphometric data. *Syst Biol*. 2010;59:245–61.
- LaBelle JM. Coffee cans and Folsom points: why we cannot continue to ignore the artifact collector. In: Zimmerman L, Vitelli K, Hollowell-Zimmerman J, editors. *Ethical Issues in archaeology*. Walnut Creek: Altamira; 2003. p. 115–27.
- Lycett SJ. Understanding ancient hominin dispersals using artefactual data: a phylogeographic analysis of Acheulean handaxes. *PLoS ONE*. 2009;4(10):e7404. doi:10.1371/journal.pone.0007404.
- MacLeod N. Phylogenetic signals in morphometric data. In: MacLeod N, Forey P, editors. *Morphology, shape and phylogeny*. London: Taylor & Francis; 2002. p. 100–38.
- Maddison WP, Maddison DR. Mesquite: a modular system for evolutionary analysis. (<http://mesquiteproject.org/mesquite/mesquite.html>). Accessed 26 May 2009.
- Mesoudi A, O'Brien MJ. The cultural transmission of great basin projectile-point technology II: an agent-based computer simulation. *Am Antiquity*. 2008;73:627–44.
- Morrow JE, Morrow TA. Geographic variation in flute projectile points: a hemispheric perspective. *Am Antiquity*. 1999;64:215–30.
- Newell ND. Fossil populations. In: Sylvester-Bradley P, editor. *The species concept in palaeontology*. London: The Systematics Society; 1956. p. 63–82.
- Nunn CL, Arnold C, Matthews L, Mulder MB. Simulating trait evolution for cross-cultural comparison. *Philos Trans R Soc B*. 2010;365:3807–19.
- O'Brien MJ, Lyman RL. *Cladistics and archaeology*. Salt Lake City: University of Utah; 2003.
- O'Higgins P, Jones N. Tools for statistical shape analysis. Hull York Medical School, <http://hyms.fme.googlepages.com/resources>. Accessed 13 October 2009; 2006.
- Polly PD. On the simulation of the evolution of morphological shape: multivariate shape under selection and drift. *Palaeontol Electronica*. 2004;7A(2):1–28. [http://palaeo-electronica.org/paleo/2004\\_2/evolution2\\_04.htm](http://palaeo-electronica.org/paleo/2004_2/evolution2_04.htm). Accessed 12 August 2010.
- Ratto N. Estrategias de Caza y Propiedades del Registro Arqueológico en la Puna de Chaschuil (Depto. De Tinogasta, Catamarca, Argentina). Unpublished Ph.D. Dissertation, Dept. of Anthropological Sciences, University of Buenos Aires; 2003.
- Sepkoski D. The emergence of paleobiology. In: Sepkoski D, Ruse M, editors. *The paleobiological revolution: essays in the growth of modern paleontology*. Chicago: University of Chicago; 2009. p. 15–42.
- Shott MJ. Spears, darts, and arrows: late woodland hunting techniques in the Upper Ohio Valley. *Am Antiquity*. 1993;58:425–43.
- Shott MJ. Transmission theory in the study of stone tools: a Midwestern North American example. In: Barton M, Clark G, editors. *Rediscovering Darwin: evolutionary theory and archeological explanation*. Archeological Papers of the American Anthropological Association No. 7. Washington, D.C.; pp. 193–204; 1997.
- Shott MJ. Darwinian evolutionary theory and lithic analysis. In: O'Brien M, editor. *Cultural transmission and archaeology: issues and case studies*. Washington, DC: Society for American Archaeology; 2008a. p. 146–57.
- Shott MJ. Lower paleolithic industries, time, and the meaning of assemblage variation. In: Holdaway S, Wandsnider L, editors. *Time in archaeology: time perspectivism 20 years later*. Salt Lake City: University of Utah; 2008b. p. 46–60.
- Shott MJ. Morphometric approaches to the study of fluted points. Paper presented at the 74th Annual Meeting of the Society for American Archaeology. Atlanta; 2009.
- Shott MJ. Crises and solutions in American archaeology. *SAA Archaeol Rec*. 2010a;10(1):37–8.
- Shott MJ. Size-dependence in assemblage measures: essentialism, materialism, and 'SHE' analysis in archaeology. *Am Antiquity*. 2010b;75:886–906.
- Shott MJ, Trail BW. Exploring new approaches to lithic analysis: laser scanning and geometric morphometrics. *Lithic Technol*. 2010;35:195–220.
- Sinopoli CM. Style in arrows: a study of an ethnographic collection from the Western United States. *Mich Discuss Anthropol*. 1991;10:63–87.
- Stern N. The structure of the lower pleistocene archaeological record: a case study from the Koobi Fora formation in Northwest Kenya. *Curr Anthropol*. 1993;34:201–25.
- Sumner TA, Riddle AT. A virtual paleolithic: assays in photogrammetric three-dimensional artifact modelling. *PaleoAnthropology*. 2008;2008:158–69.
- Tëmken I, Eldredge N. Phylogenetics and material cultural evolution. *Curr Anthropol*. 2007;48:146–53.
- Trigger B. *A history of archaeological thought*. 2nd ed. Cambridge: Cambridge University Press; 1996.
- Wilhelmsen KH. Building the framework for an evolutionary explanation of projectile point variation: an example from the Central Mississippi River Valley. In: Hunt T, Lipo C, Sterling S, editors. *Posing questions for a scientific archaeology*. Westport: Bergin & Garvey; 2001. p. 97–144.
- Williams G. *Adaptation and natural selection: a critique of some current evolutionary thought*. Princeton: Princeton University Press; 1966.
- Wilson T. Arrowpoints, spearheads, and knives of prehistoric times. *Annual Report of the United States National Museum for 1897*. Washington, D.C.; 1899. pp. 1–267.