A. L. KROEBER AND THE MEASUREMENT OF TIME'S ARROW AND TIME'S CYCLE

R. Lee Lyman and Judith L. Harpole
Department of Anthropology, University of Missouri, Columbia, MO 65211

A. L. Kroebler invented frequency seriation in 1916 to measure the passage of time. This chronometer measured the flight of time's arrow with artifact types that changed their relative frequencies in a uniform fashion over what seemed, on the basis of other evidence, to be the linear passage of time. In 1919 and again in 1940, Kroebler sought to measure cycles of cultural phenomena over time and chose a different method for doing so. This method, now termed time-series analysis, was implemented by plotting the averages of several metric dimensions of a kind of artifact against their known age. Kroebler observed recurring averages and concluded that cultural phenomena cycled through time. Frequency seriation and time-series analysis as implemented by Kroebler comprise significant differences in the units they use, their protocols, and their analytical goals.

As we enter the third millennium, Americanist archaeologists are taking frequent glances at the history of their discipline (e.g., Lyman, O'Brien, and Dunnell 1997; Lyman, Wolverton, and O'Brien 1998). Some of these glances have focused on the history of various chronometric tools (Lyman and O'Brien 1999, 2000, 2001; Nash 2000; O'Brien and Lyman 1999), perhaps because the archaeological record is a modern phenomenon, yet archaeologists wish to know when portions of that record were formed. Knowing how the age of record has been determined is therefore a necessary part of being an archaeologist. We find this justification for the historiography of our discipline sufficient but also think that it is more interesting, if not necessary, to consider the underlying epistemological and ontological bases of chronometers and related analytical tools. It seems to us that...
A. V. Kidder, Nels Nelson, and others working in the early twentieth century held some implicit notion of cultural development or evolution. We say this because at least some of the variation in the artifacts comprising the archaeological record was thought by these and other archaeologists to reflect cultural change and thus the passage of time (Lyman, O’Brien, and Dunnell 1997). We emphasize the word “change” because it not only denotes the passage of time but also implies some sort of connection or continuity between earlier and later archaeological manifestations above and beyond mere relative position in time.

Our focus on historical matters here involves explication of the distinction between two of the analytical techniques used by A. L. Kroeber to measure time’s passage and cultural change. Because these techniques are sometimes confused by anthropologists, we explore the assumptions which underpin each and by example illustrate how the measurement units Kroeber used influenced his analytical results. One analytical technique is frequency seriation, a patently archaeological invention (Lyman, O’Brien, and Dunnell 1997; Lyman, Wolverton, and O’Brien 1998; O’Brien and Lyman 1999); the other analytical technique is time-series analysis (e.g., Chatfield 1975). Kroeber invented frequency seriation to measure time’s passage; he used time-series analysis to monitor cultural cycles. As will become clear, the measurement units Kroeber used for each analytical technique are distinctly different, and the analytical protocols of each technique are equally distinct. These epistemological differences have not previously been explored, and this is perhaps why the results of the two techniques have been said to monitor the same thing (Teltser 1995). Prior to presenting our viewpoint on these matters, however, several things need to be made clear.

**Prefatory Remarks**

The historical aspects of our discussion are grounded in the perspective of evolutionary archaeology, the most important aspect of which in the present context is that cultural transmission and heritability underpin any chronometer that employs formal variation in artifacts to measure the passage of time (Lipo 2001a, 2001b; Lyman and O’Brien 1998; Neiman 1995; O’Brien and Lyman 1999, 2000). We use Gould’s (1987) metaphors of “time’s arrow” and “time’s cycle” to denote the fact that, although time is linear, it can be measured as either a nonrecurrent linear phenomenon or as a recurrent cyclical phenomenon, respectively. As we will show, whether one measures time’s arrow or time’s cycle depends on the measurement units used, such as the month of January measuring time’s cycle every solar year, whereas the solar year AD 2001 measures time’s arrow since the birth of Christ. Because of this possibility, we disagree with McGee and Warms’s (2000:142, n. 2) assessment of Kroeber as believing that time’s cycle would be found in “all manner of cultural phenomena.” We agree, however, with McGee and Warms’s (2000:142, n. 3) assessment of Kroeber’s use of time-series analysis as being underpinned by an “evolutionary perspective” because the artifacts he used in that analysis comprise an evolutionary lineage.

We define a time-series analysis as comprising the arrangement of phenomena
in a temporal sequence based on the respective ages of the phenomena and subsequent study of changes in the phenomena across the ordering (Braun 1985; Chatfield 1975). We define seriation as "a descriptive analytical technique, the purpose of which is to arrange comparable units in a single dimension (that is, along a line) such that the position of each unit reflects its similarity to other units" (Marquardt 1978:258). As Rowe (1961:326) put it, the "logical order on which the seriation is based is found in the combination of features of style or inventory which characterize the units, rather than in the external relationships of the units themselves," where "external relationships" include age. Archaeological seriation, of course, typically is done for purposes of measuring time's arrow. Thus the important point to remember is that time-series analysis has a rather different analytical goal than seriation. For the former, time is known and changes in phenomena are monitored over time; for the latter, a measure of historical or linear time is the goal, and phenomena are ordered so as to (hopefully) reflect time's passage.

We define a chronometer as a means of measuring the linear flight of time's arrow. Gould (1987:157) notes that "a chronometer of history has one, and only one, rigid requirement—something must be found that changes in a recognizable and irreversible way through time, so that each historical moment bears a distinctive signature" (emphasis added). Such chronometers are precisely what early Americanist culture historians sought (Lyman, O'Brien, and Dunnell 1997). Early chronometers they developed and used include superposition implemented by stratigraphic excavation, percentage stratigraphy, two seriation techniques, and the direct historical approach (Lyman and O'Brien 1999, 2001; O'Brien and Lyman 1999). Each of these chronometers was used successfully only when Gould's "rigid requirement" was met. If the thing(s) monitored recurred, then the chronometer was faulty because it rendered time not as the linear flight of an arrow but rather as a set of reversals, recurrences, and cycles. Time measured in the latter fashion is of no use to geologists, paleontologists, and archaeologists who are interested in the temporally linear history of things such as biological descent with modification and cultural change (Lyman and O'Brien 2000). We suspect most archaeologists are aware of the recurrence problem when they seek to build, or use, a chronometer. A familiar example is found in the conversion of radiocarbon ages to calendar years; when converted, some radiocarbon ages give more than one calendar age.

Time-series analysis is, in our view, not a chronometer in the sense indicated in the preceding paragraph. Seriation—as defined above—is often used by archaeologists as a chronometer. The distinction between time-series analysis and seriation is, however, sometimes ignored and the two methods conflated. For example, despite their own characterizations of their work as "seriation," Old World prehistorians performed time-series analyses when they plotted frequencies of types of stone tools against their ages as determined stratigraphically and interpreted the variation in those frequencies (Collins 1965; Mellars 1965). This does not mean that time-series analysis cannot be used to develop new chronometers, to discover additional chronometric units, or to refine existing
chronometric units so as to have them reflect shorter temporal increments. Binford (1962), for example, calculated a simple linear regression formula that expressed the relation between pipe-stem hole diameter and the age of the pipe. Harrington (1954) had originally documented the relation between hole diameter and age and had used thirty-year increments to group pipes of known age into sets that comprised adequate samples for determining if hole diameter in fact decreased consistently over time. Harrington found such consistency—there were no reversals toward holes of larger diameter—and concluded that hole diameter was a good chronometer because there were no cycles or recurrences over time of particular diameters. Binford (1962) desired a more precise way to estimate the absolute age of individual assemblages of pipes, and his regression formula provided a way to calculate a "mean date" for such assemblages.

Plog and Hantman (1986, 1990) followed Binford's lead and regressed the age of Southwestern ceramic types against the attributes comprising the types in an effort to develop chronological units of finer temporal resolution than were provided by the types alone. Braun (1985) regressed the thickness of ceramic vessel walls against time to study changes in that variable. O'Brien and Lyman (1999:135–136) point out a potential problem with this kind of use of time-series analysis. In particular, the regression equation comprises an empirical generalization derived from the specimens studied. Whether it applies to other specimens of the same type is a separate issue, although various data indicate it may not (e.g., Deetz and Dethlefsen 1965; Hoard 1992). In particular, archaeologists, geologists, and paleontologists measure time indirectly with formal variation in artifacts, strata, and fossils, respectively, and that formal variation can occur differentially in time across geographic space. The message here is simple: chronometers and the units they comprise must constantly be tested to insure that they are in fact chronometric, such that not only does "each historical moment bear a distinctive signature" (Gould 1987:157), but also that the signature is equally historically distinctive irrespective of where in geographic space it is used as a chronometer. Of course, once the ages of specimens have been determined with a chronometer, then time-series analysis is possible.

Because we spend considerable time discussing frequency seriation, we need to briefly describe this analytical technique. Although frequency seriation has been discussed by numerous individuals (see references in O'Brien and Lyman 1999 for an introduction and Lipo 2001a, 2001b for details), the three basic requirements of the technique are not always spelled out. We phrase them here in the typical sense of ordering assemblages of spatio-temporally associated artifacts (discrete objects owing one or more of their attributes to human activity). First, assemblages must represent relatively brief intervals of time; that is, the duration of their formation—the time between the deposition of the first specimen and the time of the deposition of the last specimen—must be short. This requirement not only insures that time's passage is measured rather than the durations of formation of assemblages (Dunnell 1970, 2000), but it also begs the question of how short the duration of the formation of an assemblage must be. We return to this question later and here simply note that evolutionary theory-(von Vaupel Klein 1994) and computer
modeling (Lipo et al. 1997; Neiman 1995) suggest that a single generation—such as the modern market economy’s fiscal year—is too short of a temporal unit for producing frequency-distribution curves that fluctuate smoothly and unimodally.

Second, the seriated assemblages must derive from the same cultural tradition, where a cultural tradition is a lineage resulting from transmission. In other words, the seriated assemblages must all derive from the same line of heritable continuity (O’Brien and Lyman 1999). The third requirement rests on the fact that transmission has both a temporal and a spatial aspect and demands that the seriated assemblages all come from the same local area. This requirement limits spatial variation in order to measure time; if one is interested in measuring transmission across space, then time must be limited (e.g., Lipo 2001a, 2001b; Lipo et al. 1997; see also Dunnell 1970, 1981; O’Brien and Lyman 1999).

The frequency-seriation model holds that historical types (1) will occur during only one span of time and (2) will have unimodal (relative) frequency distributions over time (Dunnell 1970). Originally explained as resulting from the “popularity” of types (Lyman, O’Brien, and Dunnell 1997), it is clear that this is not an explanation of what has been observed but is, rather, an empirical generalization that simply restates what has been observed (Dunnell 2000; Telser 1995). It is now clear that so-called historical types produce unimodal frequency distributions when seriated because the transmission processes that result in their persistence over time are stochastic. As Lowe and Lowe (1982:540) note, “styles are stochastic systems, neither entirely random nor deterministic, but instead a mixture of structure and chaos.” The vagaries of transmission and replication result in the battleship-shaped, or unimodal, frequency distributions of seriable historical types (if the durations of the seriated assemblages are correct), a fact confirmed multiple times in computer simulations performed by archaeologists (e.g., Lipo et al. 1997; Neiman 1995) and paleontologists (e.g., Raup and Gould 1974; Raup et al. 1973). Seriable types must be adaptively neutral; that is, the types must be functionally equivalent relative to one another so that they reflect only the vagaries of transmission. As Hole and Shaw (1967:86) put it, types that are successfully arranged via frequency seriation “are likely to have relatively neutral adaptive value in a culture.”

Dunnell (2000:549) has pointed out that when ordering assemblages via frequency seriation, the analyst could “insist on an exact match with the unimodal [frequency distribution] model before regarding the order as chronological, a deterministic solution; alternatively, one could accept the ‘best fit’ to the unimodal model as chronological, a probabilistic solution” (emphasis in original). He notes that deterministic solutions may seldom be found but does not elaborate on why this should be so. We suspect one important factor contributing to the regular failure to find deterministic solutions concerns the durations of the seriated assemblages, and we return to this issue later.

In the following we are concerned with Kroeber’s use of frequency seriation and his use of time-series analysis. Frequency seriation is a chronometer that uses variation in the frequencies of formal artifact types to measure time as a nonrecurrent linear phenomenon; time-series analysis is meant to measure change,
whether linear or cyclical, in artifacts already arrayed against the linear flight of
time's arrow. Kroeber knew the difference between the two analytical techniques
and the role that the linear passage of time played in each, even though he did not
refer to what he did as frequency seriation or as time-series analysis.

KROEBER AND FREQUENCY SERIATION

During the summer of 1915, Kroeber volunteered to help American Museum
personnel with their research in the Southwest (Wissler 1915:397). While walking
across the countryside around Zuni Pueblo, Kroeber collected pottery sherds from
the surfaces of more than a dozen sites. He noticed that some collections tended to
be dominated by "red, black, and patterned potsherds," whereas other collections
were dominated by "white" sherds that were "nearly always pale buff, pinkish, or
light gray" (Kroeber 1916b:8). He concluded that there "could be no doubt that
here, within a half hour's radius of the largest inhabited pueblo [Zuni], were
prehistoric remains of two types and two periods, as distinct as oil and water"
(Kroeber 1916b:9). Two lines of evidence suggested that he was dealing with
temporal differences in the variants of pottery he found. First, historical data
indicated that "several of the ruins were inhabited in Spanish times" (Kroeber
1916a:43), whereas other ruins were said by his informants to have been inhabited
"long ago" (Kroeber 1916b:9). Second, Kroeber (1916a:43) observed that historic
"ruins normally include standing walls, and loose rock abounds. [Prehistoric] sites
are low or flat, without walls or rock, [probably] due to the decay of age, or to the
carrying away of the broken rock to serve as material in the nearby constructions
of later ages." Based on these two lines of evidence, Kroeber (1916b:9-10)
concluded that white ware was "wholly prehistoric," whereas black and red ware
"is the more recent [belonging] in part to the time of early American history."

Kroeber (1916b) presented the absolute frequencies of ten types of sherds he
collected from what he took to be historic sites in one table and the absolute
frequencies of those same ten types of sherds from what seemed to be prehistoric
sites in a second table. In a third table, he presented the relative or proportional
frequencies of each of the ten types from all sites he had visited, including Zuni
Pueblo. His fourth table is critically important because in that table Kroeber
lumped his ten types into three more general types and presented the relative
frequencies of these general types from each site. A graph we made of the relative
frequencies of those three general types across the ordering of sites Kroeber
ultimately presents as a tentative chronology does not comprise a deterministic
frequency seriation; none of the three general types displays anything
approximating a unimodal frequency distribution (Figure 1). So far our efforts to
sort these collections on the basis of these three types have failed to more closely
approximate a deterministic solution.

Paced with such a problem and lacking computer-aided analysis techniques,
Kroeber seems to have taken the only possible, yet remarkably simple, route to a
nearly perfect deterministic solution. Kroeber's original ten types included both
black corrugated ware and white corrugated ware, but those color differences
seemed of no chronological significance. However, corrugated ware in general seemed to be relatively old, so Kroeber lumped the two colors into the general type corrugated ware. The collections of pottery from individual sites were arranged so that the relative abundance of corrugated ware decreased monotonically, with two exceptions, as indicated in his fifth table of data. Those data for corrugated ware are graphed in Figure 2. The basis for this arrangement was Kroeber's (1916b:15) impression that corrugated ware, given its rare association with modern pottery types and its regular association with decayed ruins, gradually decreased in relative frequency as time passed; this allowed him to arrange “the sites in order accordingly” and would in 1916 come to be known as the “popularity principle,” originally stated, but not named as such, by Nels Nelson (1916).

Kroeber accounted for one of the two exceptions in his ordering—Kyakki W2—by considering it to comprise a nonrepresentative sample; only twenty-five sherds had been collected from this site. The other exception—Towway—was placed in the order on the basis of one of Kroeber’s original ten types, what he called “Three Color.” The relative abundance of “Three Color” increased monotonically once it appeared in the sequence and was most abundant in the modern Zuni assemblage. Here, Kroeber was anticipating the observation made fifty years later that historical types may be missing from collections dating to the initial and the final periods of a type’s duration (Dunnell 1970). He thus used a second type to finalize the ordering of the collections. Other types were simply appended to the arrangement, and thus their frequencies tend to fluctuate without pattern. We graphed only one of these other types—black-on-red—in Figure 2, as Kroeber did not present the relevant data in a fashion conducive to graphing all
Figure 2. Centered-Bar Graph Showing the Results of Kroeber’s Final Frequency Seriation Based on the Relative Frequencies of Corrugated Ware and Three-Color Ware

Only one—black-on-red—of several other types is included for comparison to show that the two types on which Kroeber based his final seriation fluctuate unimodally. See the text for discussion of the exceptional Towway and Kyakki W assemblages. Some site names are abbreviated.

types. Kroeber suggested that at least some of the fluctuation in these other types could be attributed to sampling error.

Three important aspects of what Kroeber did warrant brief discussion. First, his lumping of types indicates that he conceived of those units as things created by the analyst, as opposed to things to be discovered. Types were units that allowed—if properly constructed by the analyst—measurement of the flight of time’s arrow. The nineteenth-century phyletic seriations (Lyman, O’Brien, and Dunnell 1997; O’Brien and Lyman 1999) of British gold coins by John Evans (1850, 1875), the phyletic seriation of Egyptian pottery by W. M. F. Petrie (1899), and other such efforts (e.g., Pitt-Rivers 1870, 1875a, 1875b; Uhle 1903) also measured that flight, but they played no role in Kroeber’s invention of frequency seriation (see also Trigger 1989:200–202). That this is so is clearly reflected in two ways. First, it is reflected by the fact that Kroeber used ideational or conceptual units to categorize his sherds; such units are required by frequency seriation but not by phyletic seriation. The presence/absence of empirical units is typical of phyletic seriation (O’Brien and Lyman 1999). Kroeber used relative frequencies of empirical specimens categorized as one of two conceptual units—corrugated ware and three-
color ware—to seriate his collections. Second, Kroeber assessed the similarity of collections in a unique way—by measuring and comparing the relative frequencies of multiple types. Phyletic seriation simply measures the similarities of specimens, not the similarities of frequencies of specimens within categories.

The second important aspect of Kroeber’s work at Zuni is that there was no explicit theoretical warrant for inferring that his seriated collections measured the passage of time, although it seems to us that he had in mind some notion of cultural change. In particular, heritable continuity affected by cultural transmission, if reflected in pottery, would also reflect (and allow the measurement of) the linear passage of time. Kroeber knew that artifacts could (and often did) change formally over time (how they might change was what was unclear); plus he had observed the association of different pottery types with ruins inferred to be of different ages. He was very aware of the hypothetical nature of the inference and noted “I have not turned a spadeful of earth in the Zonii country” (Kroeber 1916b:14), but he also stated that the “final proof [of his hypothesized chronology] is in the spade” (Kroeber 1916b:20). The latter is a classic statement in the history of Americanist archaeology because, like his contemporaries, Kroeber assumed that stratigraphically superposed collections would provide the empirical proof that his ordering of collections was in fact chronological (Lyman and O’Brien 1999). Leslie Spier (1917a, 1917b) collected that proof in 1916.

The third important aspect of Kroeber’s seminal frequency seriation resides in his (1916a:44) observation that the seriated assemblages “shade[d] into one another” and there was “no gap or marked break between” the prehistoric and historic periods. The historic and prehistoric periods might have been “as distinct as oil and water” (Kroeber 1916b:9) when distinguished on the basis of the degree of deterioration of associated ruins, but these were not the cultural phenomena used in the seriation. Kroeber (1916b:15) indicated that the historic and prehistoric periods “can normally be distinguished without the least uncertainty, and the separateness of the two is fundamental,” but—and this is significant—his frequency seriation of pottery indicated the two periods “do not represent two different migrations, nationalities, or waves of culture, but rather a steady and continuous development on the soil.” This was the clearest expression to that point in the history of Americanist archaeology of the belief that time could be measured archaeologically as a continuous linear (noncyclical) dimension because the way in which artifacts were categorized suggested ethnic and thus heritable continuity. This point was missed completely over the next twenty years as Americanist archaeologists came to adopt stratigraphic excavation as the preferred method of building chronologies of cultures (Lyman and O’Brien 1999; O’Brien and Lyman 1999).

Thirty-five years after Kroeber seriated his Zuni potsherds, he remarked that his effort was of “historical interest as regards method” for two reasons (Kroeber 1952:230). First, his work took place simultaneously with what he called the first “stirrings” of “confidence that time differences could be ascertained” in the archaeological record of the Americas (Kroeber 1952:230). Second, he had had no time to excavate and thus could not use the stratification of sediments as a
chronometer; therefore he was forced to make what he now called “stylistic seriations” that he hypothesized comprised a “temporal sequence” (Kroeber 1952;230). We add to Kroeber’s two historical reasons that his work is significant the fact that it rested implicitly on the patently evolutionary notion of heritable continuity among artifacts affected by cultural transmission (Lyman 2001); artifacts and assemblages thereof are homologously similar. Frequency seriation provides a means to measure the continuous linear flight of time’s arrow because it measures the evolution of artifact lineages.

The term “frequency seriation” was not used by Kroeber (1916a, 1916b) to describe what he did, nor did he use the term “seriation” until thirty-six years later (Kroeber 1952). The latter term was used by Spier (1917a) in reference not only to Kroeber’s (1916a, 1916b) work but also to the very different work of Kidder (1917). Spier could use the term correctly to refer to both of those efforts because although Kroeber and Kidder followed distinctly different analytical protocols (Lyman, O’Brien, and Dunnell 1997; Lyman, Wolverton, and O’Brien 1998; O’Brien and Lyman 1999), both men were ordering artifacts based on the similarities of those artifacts.

Seriation as an analytical method employed by archaeologists today comprises three different techniques, each of which involves the process of placing phenomena in an order or series based on a specific measure of similarity based in turn on a generic principle of arrangement (Dunnell 1970; O’Brien and Lyman 1999; Rouse 1967; Rowe 1961). The generic principle involves formal similarity; the more similar two phenomena are formally—in terms of their attributes—the closer together they are placed in the arrangement. Three specific measures of similarity have been used by archaeologists, only one of which comprises frequency seriation. First, similarity can be assessed by noting which attributes of discrete objects are shared among them. This was the protocol of phyletic seriation followed by Kidder (1917) when he mimicked what Petrie (1899) and others before him had done. Second, similarity can be assessed by noting which types are shared by multiple assemblages of artifacts; this is known as occurrence seriation (Dunnell 1970), a technique invented in 1963 (Dempsey and Baumhoff 1963). The third technique for assessing similarity is the one Kroeber invented.

When Kroeber (1916a, 1916b) invented frequency seriation, he did two important things above and beyond using the generic ordering procedure of seriation. First, he assessed similarity in a new way—he noted the relative frequency of a kind of artifact in each of multiple collections of artifacts and then ordered those collections based on the similarities of the relative frequencies of that kind of artifact. This is what made his analytical protocol not only innovative but also distinct from that of Kidder and others who had used a protocol like Kidder’s in the nineteenth century. Second, like those archaeologists who preceded him in using seriation as an ordering technique, Kroeber inferred that the resulting order was chronological; that is, he interpreted the resulting order as one that measured the passage of time.

Kroeber (1909;5) had earlier noted not only that a stratigraphically ordered set of artifacts described by Uhle (1907) was in fact chronological but also that Uhle’s
ordering showed little more than “passing change in fashion.” Such “stylistic pulsations,” as Wissler (1916:195–96) and Spier (1918:201) referred to them, would shortly come to occupy some of Kroeber’s analytical energy. Importantly, Kroeber’s (1916a, 1916b) discussion makes it clear that he was trying to measure the linear flight of time’s arrow by monitoring changes in the relative frequencies of variant kinds of artifacts. After several unsuccessful attempts to order his collections based on the relative frequencies of particular types, Kroeber succeeded in defining two types that displayed unimodal frequency distributions over what seemed to be the passage of time based on the condition of the ruins with which the sherds were associated. The significance of this point regarding the influence of artifact typology on analytical results cannot be overemphasized.

To produce a frequency seriation of types, the frequencies of which fluctuate unimodally and thus theoretically measure the flight of time’s arrow, the types, or analytical units, must be defined in particular ways. This is what Kroeber (1916a, 1916b) grappled with and what makes his work of significantly more importance than mere historical interest. The types used in a frequency seriation must comprise what came to be variously termed “styles” or “historical types”; in particular, their frequencies must comprise, in Gould’s (1987) words, historically distinctive signatures. The relatively smooth unimodal frequency distributions of the types Kroeber ultimately defined (see Figure 2) suggested to him that he was measuring not only time’s passage but “a steady and continuous development” of an artifact lineage (Kroeber 1916b:15). The reason that frequency seriation worked as a chronometer was because it measured what later was referred to as “historical continuity” but what is better referred to as “heritable continuity” (O’Brien and Lyman 1999). Frequency seriation can measure cultural transmission, an evolutionary process that takes place over time (Lipo 2001a, 2001b; Lipo et al. 1997; Neiman 1995; Teltscher 1995). Certainly, types might be identified that change in relative frequency monotonically over time but that do not measure heritable continuity, but we suspect this will be rare, if not improbable.

A DIFFERENT KIND OF ANALYSIS

Shortly after his work at Zuni, Kroeber used, if implicitly, the heritable-continuity notion, but he employed an entirely different kind of analytical unit and a different analytical protocol than he had used in his seminal frequency seriation. The unit Kroeber used a few years after his work with pottery sherds from the Southwest was similar, but not identical to, the sort of unit originally used by Kidder (1917). In his famous study of changes in women’s fashions, Kroeber (1919:235) sought to formulate a law concerning the “growths and declines [of] normally recurring events.” He reasoned that a “generic [lawlike] principle” regarding the “rhythmic inevitability” of change in cultural phenomena could be found only with “verifiable measurements,” which, when plotted against time, would display “a polygon of frequency or normal curve such as the statistical sciences employ” (Kroeber 1919:236, 237). But here Kroeber was not searching for the unimodal curve given by a type’s frequency in a frequency seriation;
instead, each curve he now sought would be multimodal. Further, the curves or lines would here monitor changes in the observed averages of "measurements" of phenomena—each average representing a particular year—rather than the frequencies of various kinds of measurements such as types of sherds. Finally, the measurements and the lines they defined were to be arrayed against time's arrow, making Kroeber's study of women's fashions a time-series analysis. This was significantly different than what he had done with Southwestern pottery a few years earlier. When he invented frequency seriation, the desired inference was the linear flight of time's arrow; when he measured women's fashions, Kroeber knew the flight of time's arrow and wished to infer patterns of cultural change, specifically cycles of cultural change rendered as multimodal curves, not the flight of time's arrow from the unimodal curves of frequency seriation.

For his study of women's fashions, Kroeber (1919:238–39) reasoned that the measurements would have to be taken not from "utilitarian pieces [because these] do not modify freely"; instead, they must come from artifacts that do not "vary in purpose" and are "purely stylistic." Therefore, to search for the rhythmic pulses, Kroeber (1919:239) chose "women's full evening toilette [because it] has served the same definite occasions for more than a century." The function of women's evening clothing had not changed over time, but its style had; and thus "styles" of cultural phenomena, because they were adaptively neutral (Dunnell 1978; Hole and Shaw 1967), were the unit of choice for measuring cultural patterns manifest as recurrent phenomena. With respect to using averages of dimensions of women's evening wear such as dress lengths, Kroeber correctly chose units that were functionally equivalent and reflected the vagaries of cultural transmission. Those units showed what he wanted them to show—cycles in cultural phenomena—and thus they are decidedly inappropriate to frequency seriation, or any seriation technique meant to measure the linear flight of time's arrow. Such units make poor chronometric units because they fail to meet the rigid requirement of comprising a historically distinct signature (see below).

Kroeber (1919) measured eight variables on about ten illustrations per year as shown in the pages of various magazines published between 1844 and 1919 inclusively. He collaborated with Jane Richardson twenty years later, increasing the temporal range included in his sample (Richardson and Kroeber 1940; see also Kroeber 1948:331–36; Kroeber 1957). Kroeber (1919) and Richardson and Kroeber (1940) plotted the average of each measured variable against time to produce a line that fluctuated as time passed; the shape of the various lines was interpreted to reflect the process and pattern of stylistic change in women's fashion. Although Kroeber knew that he was monitoring heritable continuity, he assumed that he was using adaptively neutral units—"styles"—to do so. Lowe and Lowe's (1982) reanalysis of Richardson and Kroeber's data suggest his assumption was correct. With respect to how he dealt with those units analytically, however, Kroeber clearly was not following the procedure of frequency seriation he had earlier invented.

Kroeber did not classify fashions as to type, nor did he plot the frequencies of those types against time. For his frequency seriation of Southwestern pottery
KROEBER: MEASUREMENT OF TIME'S ARROW AND TIME'S CYCLE 325

sherd s, Kroeber used conceptual units known as types. For his study of women's evening wear, Kroeber used units expressed as averages of dress length and other dimensions. The latter analysis revealed cycles, or repetitions, in the occurrence of the average of individual dimensions of women's fashions. These averages cannot be used as chronometers because various of them recur at different points in time. For example, average dress length per decade both rises and falls, but only so far in either direction. Reversals in dress length create cycles, such as in Figure 3, where it is shown that, based on Richardson and Kroeber's (1940) data, between 1710 and 1920, dress length (measured from the woman's mouth to the bottom of the hem) was 95 percent of a woman's height (measured from the woman's mouth to the bottom of her foot) five different times.

In both the case of Southwest ern pottery and the case of women's evening wear, Kroeber's results were dependent on the measurement units he chose. In both cases the units reflected heritable continuity and the vagaries of cultural transmission. Some investigators have failed to note this aspect of Kroeber's work and have attempted to explain observed periodicity in various phenomena by calling upon social forces, economics, and other cultural factors as the causes of those cycles (e.g., Robinson 1975, 1976; Weeden 1977). Although such factors may in fact be causal, the linkages seem to us to be weak, just as they seemed to

![Graph showing dress length as a percentage of model height over time](image)

**Figure 3. Grand Average per Decade of Annual Averages of Dress Length in Women's Evening Wear**

The closer the broken-stick line to 100 percent, the longer the dress length. Note that if dress length is 95 percent of the height of the model (measured from the mouth of the model to the bottom of her foot), this category of dress length would not serve chronometric purposes because it occurs five times during the temporal period included in the graph. Based on data in Richardson and Kroeber 1940.)
Kroeber (1952, 1957). In this respect, we find stronger linkages and greater explanatory power elsewhere. In particular, detailed analysis suggests that at least some of the units displaying recurrences over time are not adaptively neutral (e.g., Braun 1985; Rands 1961)—that is, they are functional (Dunnell 1978)—and thus their cycles or periodicity are predictable from evolutionary theory and a knowledge of the selective environment. At present, we are not sure that the dimensions of women's evening dress measured by Kroeber are functional attributes, but further consideration of this issue is beyond our scope here.

Some later investigators interested in aspects of fashion history merely mimicked the analytical protocol Kroeber had used to study women's dresses (e.g., Robinson 1975, 1976; Weeden 1977). One analyst interested in the history of women's hats, however, attempted to mimic Kroeber's protocol of frequency seriation. Although there is nothing inherently wrong with the protocol followed, that analysis underscores how easily researchers might confuse frequency seriation as a chronometer with time-series analysis as a technique for studying cultural lineages. That some have misunderstood Kroeber's (1919) work is exemplified by misstatements of what he actually did and how it relates to frequency seriation (see next section). It is in part with the intent of precluding additional misunderstanding of that work that we offer the following discussion.

WOMEN'S BONNETS

Inspired by Richardson and Kroeber's (1940) study of temporal changes in women's dresses, Sarah Turnbaugh (1979) undertook a study of the history of women's headgear. Rather than mimic Kroeber's (1919) protocol for the analysis of fashion, Turnbaugh used what she called the "seriational technique," characterized by her as "a new method for measuring ... the incessant nature of stylistic change in fashion" (Turnbaugh 1979:242, 241). Her analytical protocol comprises two steps. The first is "content analysis [which] involves the controlled observation and systematic counting of the frequency of occurrence of symbols or traits" (Turnbaugh 1979:242). The data source comprised "the total population of illustrations" of bonnets deemed fashionable between 1831 and 1895 and illustrated in Godey's Lady's Book and Magazine (Turnbaugh 1979:243), a monthly magazine published between 1830 and 1898 in 137 volumes, 2 volumes per year (Finley 1931). Turnbaugh (1979:243, 244) identified "four major categories" of bonnet (pamela, capote, bibi, and crownless bonnet) comprising twenty-one "distinct varieties"; these were her analytical units. For each year between 1831 and 1895 inclusively, she tallied the absolute frequency of illustrations of each of the twenty-one varieties, plus three additional categories labeled "hats," "caps," and "other/nothing."

The second step of Turnbaugh's analytical protocol involves what she called a "variant of seriation"; she (incorrectly) defines seriation as "the chronological ordering of the frequency and life span of a specific trait, or an object" and also states (incorrectly) that the seriation method "necessarily assumes that the popularity of any object is transient and may be measured" (Turnbaugh 1979:243).
She continues: "After data have been collected by using content analysis, fashion counting, or other appropriate counting methods, seriation may be used to order the data for temporal analysis and interpretation" (Turnbaugh 1979:243). Recalling that seriation in general has as its goal the ordering of phenomena based on some principle of similarity, and that it is only inferentially a chronometer, what Turnbaugh did was to use similarity in time as the principle of ordering. Turnbaugh already knew the chronological order of the bonnets given the publication dates of the magazine issues, and she simply plotted the frequencies of illustration of each variety of bonnet against the year when those illustrations appeared. The direction of flight of time's arrow was already known, and the frequencies of kinds of headgear were simply plotted against that flight. Turnbaugh's analysis, like Kroeber's concerning dress lengths, is properly referred to as a time-series analysis rather than a seriation, although it is not unusual to find archaeologists labeling as seriation virtually any technique for plotting frequencies of artifacts against time (e.g., Deetz and Dethlefsen 1965; see also Lyman, Wolverton, and O'Brien 1998).

Turnbaugh's (1979:243) statement that seriation was an archaeological invention "developed as an analytical tool for the study of prehistoric cultural change" is a mischaracterization of history. In our view, all three seriation techniques were developed as chronometers that rest on the assumption that material culture changes over time. Turnbaugh (1979:243) is correct that frequency seriation "creates a type of bar graph" but is incorrect that the "frequency of occurrence is plotted against a time line to produce a 'battleship curve'." Recall that time is inferred from the ordering resulting from the frequency seriation and that the ordering seeks to approximate or mimic a unimodal curve; in time-series analysis, the phenomena plotted against time may not produce a battleship-shaped curve, and it is the shape of the curve that is of analytical interest. A redrawn version of Turnbaugh's bar graph showing what she characterizes as the "stylistic evolution of American [women's] headgear" (Turnbaugh 1979:243) is given in Figure 4. We note that if we adopt the probabilistic solution of frequency seriation, then some, but not all, of Turnbaugh's twenty-one varieties of bonnets approximate battleship-shaped curves. In Figure 4 and our analyses, we eliminated her "hat," "cap," and "other/nothing" headgear categories—A, B, and C, respectively. These types would be of little utility in a chronometer seeking an ordering within the time period 1831 to 1895 because they span the entire period.

Turnbaugh (1979:243) cites archaeological references to seriation (e.g., Dunnell 1970), but she does not cite Kroeber's (1916a, 1916b) seminal effort or Spier's (1917a, 1917b) follow-up study. That she, in fact, did not perform a frequency seriation in the sense Kroeber intended when he invented the technique is clear from the facts that in Turnbaugh's case (1) time's direction was known and served as the basis for the ordering, (2) absolute frequencies rather than relative frequencies of varieties of women's headgear per year were plotted, and (3) absolute frequencies of illustrations rather than relative frequencies of the artifacts themselves were used (see Figure 4). Of course, one could argue that illustrations of artifacts are themselves artifacts that might serve as the units in a frequency seriation or as the units in a time-series analysis. To use the ordering of illustrations
of artifacts as a basis for determining the age of an artifact itself rather than an illustration of it, however, requires the assumption that there is a direct correlation between how often a style is illustrated and how many individual artifacts of that style were manufactured. That this assumption is perhaps at least partially met by Turnbaugh’s sample is suggested by the fact that she was able to correctly date three of four bonnet specimens of known age using her temporal ordering, and the fourth bonnet was dated within ten years of its true age (Turnbaugh 1979:246–47). This test was successful because Turnbaugh (1979:247) established what she variously termed the “date range” or “seriational date” of each bonnet variety based on her time-series analysis (see Figure 4). Her test also suggests that the varieties of bonnet she used in her analysis may at least approximately comprise historical types in the sense Kroeber intended when he was dealing with Southwestern pottery sherds.

Questions remain, however. Are the bonnet varieties seriable—are they good chronometers—in Kroeber’s original sense? Because the temporal order of the bonnets is already known, we rephrased this question: Do the relative frequencies of bonnet varieties approximate unimodal curves when plotted against time? Further, is a year a good temporal unit in the sense that it does not influence the shape of the frequency distribution? In fact, what should the duration of the temporal units represented by seriated assemblages be? Recall that the frequency-
seriation technique requires that the assemblages be of short duration, but we wonder how short? For example, in their classic studies of New England gravestones, Deetz and Dethlefsen (1965, 1967, 1971; Dethlefsen and Deetz 1966) used a time-series approach and chose (for unspecified reasons) to plot relative frequencies of styles of gravestones per decade; the resulting frequency distributions are good approximations of smooth, unimodal curves. Similarly, for unspecified reasons, Turnbaugh plotted the data she compiled by year, despite the fact that the magazines—the source of her data—were published monthly and would have thus allowed finer scale time-series analysis. Consideration of these various issues via reanalysis of Turnbaugh's data reveals some of the significant nuances of the frequency-seriation technique.

**REANALYSIS OF BONNET FREQUENCIES**

Few of the bonnet varieties in Turnbaugh's graph display the battleship-shaped frequency-distribution curves of a deterministic solution (see Figure 4). If we consider those varieties with more than two annual samples, only varieties H, K, P, and Q comprise deterministic solutions. Because we believe that it is indisputable that heritable continuity is intrinsic to the data graphed, there are two probable reasons why the majority of the variety-specific curves are not unimodal. First, the frequencies are absolute rather than relative. Second, each annual sample of bonnet illustrations is plotted. Simulations (Lipo et al. 1997; Neiman 1995) and theory (Teltscher 1995; von Vaupel Klein 1994) suggest that approximately battleship-shaped frequency distributions will be produced but that their edges will be jagged to greater or lesser degrees when either of these conditions hold in a frequency seriation (assuming that the jaggedness is not a result of sampling). To determine which of these conditions is influencing Turnbaugh's graph, we (1) calculated the relative frequencies of each of the twenty-one varieties and (2) generated a centered-bar graph of those frequencies plotted against time. Virtually none of the resulting curves is unimodal, as is illustrated in Figure 5, which shows a portion of the graph of relative frequencies of bonnet illustrations. This suggests either that the analytical units do not comprise good historical types in the sense of their being useful in the context of deterministic frequency seriation and/or that the annual durations of the assemblages of illustrations are inappropriate.

Recalling that in his seminal frequency seriation Kroeber (1916a, 1916b) treated his units as analytical constructs and that he lumped various types with red color into a general class he termed "Any Red" because the former, more discriminating types did not produce unimodal frequency distributions, whereas the latter, more general type did, we followed suit and simplified Turnbaugh's twenty-one varieties of bonnets. Turnbaugh (1979:244) noted that the twenty-one varieties actually comprised "four major bonnet types: the pamelia [D − F, in Figure 4], the capote [G − J], the bibi [K − T], and the crownless bonnet [U − X]" (Figure 6). Because frequency seriation is based on relative (proportional) frequencies of types and thus is subject to the problems of closed arrays (each row must sum to 100 percent), more than two types are necessary when one performs a frequency
seriation, a fact recognized by Spier (1917a). McNutt (1973, 2000) incorrectly discounts the validity of the frequency-seriation technique because he fails to grasp the implications of this point, arguing that because problems arise when only two types are used, those same problems afflict cases when more than two types are used.

Turnbaugh's four "types" comprise a sufficient number for our purposes. We determined the relative frequencies of these four types and plotted them against each year. To simplify the graph, we plotted only one row of bars for multiple contiguous years every time the relative frequencies for all types were identical over contiguous years. The resulting graph is shown in Figure 7. The frequency distributions of each type fluctuate greatly, but to us they appear to fluctuate less than in Figure 4 where each year is plotted individually. This suggests that the four basic types may serve as good historical types—they might display smooth unimodal frequency distributions—if they were not plotted by year but rather by temporal units, each of which includes multiple years.

It is necessary here to briefly consider two terms that will be of use in the following discussion. The term generation is a familiar one that is typically defined as the average time between the births of parents and the births of their offspring, calculated over multiple parents and multiple offspring (Art 1993). A perhaps less familiar term is cohort, which is defined as the group of animals born during the same time span, typically but not always a year, such that the group comprises individuals of the same age (Art 1993). Despite the fact that many (but not all) animals reproduce continuously during a solar year, both concepts are useful. The concept of a generation is useful for monitoring the history of the age of parents at reproduction and the like, and the concept of a cohort is useful for studying various demographic and mortality patterns of populations (Caughley 1977). These two
Figure 6. Four "Types" of Women's Bonnets

A, pamela; B, capote; C, bibi; D, crownless bonnet. Redrawn from Turnbaugh 1979.

Figure 7. Relative Frequencies of Types of Women's Bonnet between 1831 and 1895

Note that adjacent years with identical frequencies of all types are plotted only once. For example, the bar denoting 1895 actually comprises 1870–1895.
concepts are also useful for thinking about the duration of the formation of an artifact assemblage. Like animals, artifacts may be produced more or less continuously throughout a year, thus rendering difficult the clear distinction of one generation or one cohort from the next. The important point here is that simulations of biological reproduction and artifact replication tend to construe each temporally distinct sample of organisms or artifacts as a cohort that is parental to the temporally subsequent offspring cohort (e.g., von Vaupel Klein 1994; Neiman 1995). This has the effect of making each row of bars in a bar graph of frequencies (such as in Figures 2 and 5) equivalent to a generation in duration.

Given that the market economy rests on a fiscal year, many kinds of modern artifacts tend to have generations that are one year in duration. Such artifacts include dress fashions, automobiles, and some computer hardware and software. Other artifacts have longer generations; still others have shorter generations. We do not know the duration in years of a generation of prehistoric pottery, but we suspect that a generation of nineteenth-century bonnets approximated a year given the market economy of the time. The point to recognize is that the computer simulations performed to date suggest that, however long its duration might be for a particular category of artifact, a single generation seems to be an inappropriate duration for assemblages (cohorts) that one wishes to use in a frequency seriation. Assemblages representing longer durations comprising multiple generations would, we think, remove the jagged edges such as are apparent in Figure 5.

Because the temporal span of Turnbaugh's sample of illustrations is sixty-five years, we initially chose to divide by six to derive ten assemblages of six-years duration each and a final, eleventh assemblage spanning five years. We calculated the relative frequencies of each of the four bonnet types for those multiyear temporal units and plotted them against time. The result is more than gratifying (Figure 8), and because of this we did not perform any other permutations of lumping multiple years. With one small exception, each type displays a perfectly—non-jagged-edged—unimodal frequency distribution. The single exception in Figure 8 is that the pamela type decreases from a relative frequency of 19 percent in 1837–1842 to 2 percent in 1843–1848, and then it increases to 3.5 percent in 1849–1854. This change occurs at the end of the life span (if you will) of this type, and thus it was rarely and discontinuously illustrated (see Figure 4). It is at precisely such times, when a type rarely occurs, that not only a violation of the continuity of a historical type is permissible if not expected, but so too is it permissible (and expected) for a type's frequency to not decrease (or increase) monotonically shortly after it first appears or shortly before it disappears (Dunnell 1970). These are precisely the times when a type will be rare, and thus its frequencies in a collection will be subject to sampling error. This is in part why the previously mentioned deterministic model of frequency seriation may seldom be met.

We adopt the deterministic model of frequency seriation defined earlier. Given the preceding consideration of sampling error, we believe the fluctuation in the relative frequency of the pamela type to be insignificant for chronometric purposes. The four general types of women's headgear comprise good historical
types in the sense that their relative frequencies form the expected battleship-shaped frequency distributions over time if, and this is an important if, they are not plotted on a year-by-year—what we take to be a generation-by-generation—basis. When plotted on such a (here, annual) basis, the types do not form the unimodal frequency distributions demanded by the deterministic model, and thus their chronological utility is compromised. This returns us, then, to the question: What is a generation in the archaeological record? What comprises such in terms of the artifacts we might try to seriate? Perhaps it is those comprising a so-called occupation, but we doubt this because the life span of an artifact varies based on numerous variables such as what it is made of, its use-life history, and its potential for being recycled for the same or a different function. In our view, the issue of assemblage duration requires additional consideration, a point made previously in a different context for somewhat different reasons (Dunnell 1995; see also Lipo 2001a).

**DISCUSSION AND CONCLUSION**

Seriation is the process of ordering phenomena based on the generic principle of similarity of the phenomena; the most similar phenomena are placed close together in the ordering, the least similar are placed far apart. The ordering process has as its goal the production of the best, most consistent order possible. When used in archaeology, seriation has the second goal of building chronologies of phenomena, usually artifact types. Both goals were sought by Kroeber (1916a, 1916b) when he invented and first used frequency seriation over eighty years ago. When the method was adopted by European (and many American) archaeologists, both of these goals were met by use of the stratigraphic provenience of the artifacts.
Thus, the order of the artifacts through time was known, and the seminal European
efforts to order artifacts chronologically (Collins 1965; Mellars 1965) more closely
approximate Turnbaugh's time-series analysis than Kroeber's frequency seriation
(see also Lyman, Wolverton, and O'Brien 1998). As with Turnbaugh's analysis,
there is nothing inherently wrong with such research; the results show the history
of the frequencies of the artifact classes graphed. They are not frequency seriations
in the original sense of Kroeber; rather, they literally comprise what we term
"percentage stratigraphy" (Lyman, Wolverton, and O'Brien 1998; O'Brien and
Lyman 1999), what we take to be a particular form of time-series analysis.

The importance of classification to frequency seriation was implied by
Kroeber's (1916b) invention of the technique, and it has been noted by more recent
workers (e.g., Dunnell 1970; Lipo et al. 1997; Telser 1995). What has not been
sufficiently emphasized in these efforts is the duration of each assemblage
included in the seriation. What has been said is that all assemblages must be of
similar duration (Dunnell 1970). Here, we have underscored the important point
that each assemblage must also comprise more than the archaeological equivalent
of a single generation, whatever that might be. This helps explain why the
deterministic model may seldom be met—assemblages seriated are of
inappropriate duration. Future empirical work should mimic our efforts with
different kinds of artifacts but utilize explicitly classified artifacts. Tumbaugh
(1979) does not provide the definitive criteria of her four general types or of her
twenty-one varieties of nineteenth-century women's headgear. As a result, we
cannot explore which attributes of that headgear might be useful as chronometric
tools, and thus we cannot further explore the implications of our results with her
data set.

A. L. Kroeber was, in our view, more than clever. When he invented frequency
seriation, he gave archaeologists an analytical technique for measuring the flight of
time's arrow and the evolution of artifact lineages that depended on particular
kinds of measurement units. When he turned a few years later to a search for laws
of cultural growth and development, he employed time-series analysis, a different
analytical technique that depended on different measurement units that allowed
him to monitor time's cycle. Each analytical technique and each kind of unit were
meant to attain a particular analytical goal. Keeping the distinction between those
two techniques and two kinds of units clear is mandatory to successful
archaeological research precisely because they comprise different sorts of
analytical tools. Frequency seriation is a chronometer that uses change in the
frequencies of formal artifact types to measure time as a nonrecurrent linear
phenomenon; time-series analysis is meant to measure change, whether linear or
cyclical, in artifacts already arrayed against time. The critical analytical step
resides in the units chosen, and Kroeber recognized this.
NOTES

1. We thank C. Beck, G. T. Jones, R. L. Leonard, M. J. O'Brien, and JAR reviewers for comments on early drafts and D. Glover for advice on preparation of the figures. One portion of this paper was presented at the request of James Truncer during the Sixty-fifth Annual Meeting of the Society for American Archaeology; another portion of it was presented at the request of Todd VanPool and Marcel Harmon during the Thirtieth Annual Chacmool Conference; Mike O'Brien asked us to present a summary of the major points at a Missouri Archaeological Society Conference. We thank all these individuals for providing the catalysts that resulted in the final product.

2. Some site names are abbreviated in the text and figures for sake of simplicity.

3. Turnbaugh (1979) presented data for the years 1830–1898 in the form of a centered-bar graph like Figure 4, but she did not provide counts of observations. We used the scale she published to extract counts from her graph and note that only counts for “hats” exist for 1830, only counts for “bats” and “other/nothing” exist for 1896 and 1897, and no illustrations were published in 1898.

4. We are unaware of a statistical test to confirm or refute our impression.

REFERENCES CITED


KROEBER: MEASUREMENT OF TIME'S ARROW AND TIME'S CYCLE 337


