Whither Pattern: Just a Phase?

Kit W. Wesler
Department of Geosciences
Murray State University

Abstract

When Stanley South introduced assemblage patterns to historical archaeology in 1977, he had a quick impact on quantification and the organization of reports. However, a decade later, South himself asked, “Whither pattern?” The application of pattern studies had become rote, static, and unimaginative. Pattern studies share some fundamental principles with earlier systematic concepts in American archaeology, particularly the concept of phase. This paper explores the relationships between South’s pattern and Willey and Phillips’s phase, and suggests an expansion of the Brainerd-Robinson coefficient of similarity as one route to explore assemblage variability and comparability. I will use data from historic sites in Western Kentucky and Southeast Missouri, as well as South’s original pattern data, as examples. I emphatically do not suggest that historical archaeologists should begin defining phases. Rather, I propose that we pursue a more dynamic exploration of assemblage variability.

Presented to the Society for Historical Archaeology, Providence RI, January 17, 2003.
It’s been just over 25 years since Stanley South’s book “Method and Theory in Historical Archaeology” came out. It followed nineteen years after the publication of Willey and Phillips’s “Method and Theory in American Archaeology,” which was extremely influential to Eastern North American prehistorians’ thinking about the archaeological record. South’s book also had a great influence, especially in the way many archaeological reports have been organized and their data reported.

I’ve always wondered whether South deliberately echoed Willey and Phillips’s title. I like to think so, but I’ve never gotten around to asking him. He does not say so in his text. He cites Willey and Phillips mainly in his utilization of the horizon concept in the methodological justification for the Mean Ceramic Dating technique. Recently, I’ve been thinking more about the pattern recognition technique that South introduced in 1977, and how it relates to Willey and Phillips’s concept of phase.

South’s pattern technique was based on a functional categorization of assemblages into artifact groups: Kitchen, Architecture, Furniture, Arms, Clothing, Personal, Tobacco Pipes, and Activities (and sometimes Faunal but faunal remains are subject to a different set of preservation processes so we often leave them out). To quote South, the "basic assumption is
that each household... represents a system within a much larger system of complex variables, with the larger system imposing on each household a degree of uniformity in the relationships among its behavioral parts." "The basic postulate here is the assumption that there was a patterned casting off of behavioral by-products around an occupation site that might be viewed as a per capita, per year contribution to the archaeological record" (South 1977:86-87). South showed that the percentages of artifacts in each of his artifact groups were fairly regular across a number of eighteenth-century British colonial sites, and suggested that there was a pattern.

South's initial hope was to differentiate sites in settled areas and on frontiers, and eventually, with a great deal of data, to find ways of distinguishing cultural patterns such as British-American from German-American from French-American colonial settlements.

The patterning technique never really took hold for creative analysis. It was extremely influential as a call for quantification and for systematic collection and reporting, and also as an organizing device for describing and summarizing artifact collections.

But as a number of colleagues, including South, have noted, the application of pattern stalled pretty quickly. Orser (1989), for instance, suggested that South's patterns are not
useful because they’re too broad and are diachronically insensitive. South (1988:25), indeed, found that pattern recognition as usually practiced had been "particularistic, inductivistic exercises in identification and labeling."

I think that South hoped that historical archaeologists would explore variability in these patterns through time, space and society. Instead, we tended to use them as static standards. Either X assemblage fit the Carolina pattern, or the Architecture pattern, or it didn’t. If it didn’t, all too often the archaeologist would define a new pattern, say the 19th Century Mining pattern or the California Brothel pattern or some such. Fortunately the proliferation of named patterns didn’t last long. Unfortunately, most of us just ended up saying, so what? rather than finding ways to explore variability. I did point out at Whitehaven that there seemed to be a transition from Architecture to Carolina pattern with time and with proximity to the house, but in the absence of a series of controlled contexts representing relatively tight time spans, there wasn’t much more I could do with it. So, whither pattern? in Stan South’s words.

On the face of it, none of this sounds very much like Willey and Phillips’ concept of phase. 4 They defined the term “phase” as “an archaeological unit possessing traits sufficiently characteristic to distinguish it from all other
units similarly conceived... spatially limited to the order of magnitude of a locality or region and chronologically limited to a relatively brief interval of time.” Built into this definition was a classic formulation of the three dimensions of archaeological units: space, time, and form or content.

Theoretically, archaeological units can be based primarily on any of the three dimensions. Periods emphasize the chronological dimension, and hang spatial and cultural elements on the calendar. Willey and Phillips did not propose a taxonomy of temporal units. They did propose a taxonomy of spatial units, from site, the smallest, to locality, to region, to area. These units have not been used much.

Willey and Phillips’s most comprehensive discussion focused on the formal or content units: first component, then phase ... but then jumps to “maximum units.” There is no fully-realized taxonomy: the phase is the working unit that matters to them.

How does this relate to South’s pattern? First, because both concepts are based on the definition of archaeological content units. Second, because both sets of defined content units have been used as standards, rather than as opportunities to explore assemblage variation in creative ways. Third, because the utilization of both units has avoided tackling a statistical problem: how do we measure significant similarities
and differences among assemblages that are expressed in percentages?

In South’s patterns, we had two opposed sets of percentage ranges, with no obvious way to assess whether other assemblages were significantly like or unlike the type data, and no obvious way to explore where significant differences lay. South defined ranges, but how much deviation really is significant between two assemblages?

We have the same problem with phases. Michael O’Brien and Greg Fox have tackled this conundrum in several papers recently. In my view, what they are saying is that phases in Southeast Missouri, in particular, have been defined somewhat loosely, without regular criteria to distinguish one phase from another. They’re working around the problem of defining similarity among archaeological assemblages: how similar do assemblages have to be to belong to a phase or a pattern? How different is an outlier? How do you measure similarity?

I started thinking about whether we could use the pattern and phase data in more rigorous, that is statistical, ways. I even went to the trouble of auditing a class in multivariate statistics. What I found was that most of the standard statistical approaches to social science data didn’t fit my needs very well. To use the raw data requires fairly consistent sample sizes, which we rarely have in archaeological
assemblages. To allow for sample size biases, we either use percentages or do some fairly involved, and I think dubious, statistical manipulation. But the statistical techniques don’t do very well with percentage data, for the most part.

There are of course a number of papers that approach the problem of “Quantifying Diversity in Archaeology,” to quote the title of a 1989 book edited by Leonard and Jones. There are various techniques, borrowed from ecology, that measure something the archaeologists who use them call heterogeneity. Measures of heterogeneity combine measures of richness, having to do with the number of categories in the assemblage, and evenness, which has to do with to what extent the categories are equally represented within the assemblage. Evenness is a horror in archaeological pattern studies, where a few artifact categories—like Mississippian plain wares, or the Kitchen group—tend to dominate assemblages and the real action is in small variations in low-count, low-percentage categories.

To these “concepts related to diversity” George Cowgill also added the notions of range, standardization, and uniformity of standardization. All of this can get pretty involved statistically, but I tend to agree with Dunnell, who called these a set of measures in search of an application. For one thing, they haven’t been applied to anything that really makes them look useful. More important for the phase problem, they
measure heterogeneity within assemblages, not among assemblages. What we want to measure is similarity between assemblages, and homogeneity or heterogeneity within groups of assemblages.

O’Brien and Fox turned to the Brainerd-Robinson coefficient of similarity, which was derived specifically as a measure of similarity between assemblages where artifact counts are expressed as percentages. It measures on a scale of 200, with 200 signifying identical assemblages and 0 signifying completely different assemblages.

The Brainerd-Robinson coefficient was defined some time ago, but hit a wall in terms of application. It was proposed as a seriation device, but there are easier ways to do seriation, and we tend to think of seriation as an application to chronology. Seriation may reflect other dimensions than time, and a coefficient of similarity doesn’t specify what dimension of time, space or assemblage content is being measured, but the coefficient was typecast as a seriation device.

So Brainerd-Robinson coefficients are a technique in search of an application also, something that a few people like George Cowgill and Albert Spaulding seem to dust off now and then for heuristic purposes but haven’t applied very usefully.

In considering phase definition, O’Brien and Fox applied Brainerd-Robinson to a new purpose: showing the wide variability of collections from the same site, to make a point about sample
biases due to different collection techniques. Here I can point out that historical archaeologists noted patterned differences between surface collected and excavated assemblages a couple of decades ago.

O’Brien and Fox also used Brainerd-Robinson to measure relationships between assemblages in order to define groups of more- and less-like assemblages, that is, potentially, phases. I like the approach, but I have two problems: first that, like just about everybody, they use only ceramic type frequencies, and second, that they are fairly arbitrary about what the threshold coefficient of similarity is. I don’t see a discussion of what’s a significant level of similarity. I think there’s a further step we can take towards defining significant similarity. If we have comparable data collection techniques, we can use Brainerd-Robinson to measure assemblage similarity. We can then use very simple statistics to look at the distributions of similarity coefficients among groups of assemblages: this becomes an index of homogeneity, and a way to assess whether a single or a group of assemblages is different from others.

I constructed Brainerd-Robinson matrices for the artifact group assemblages of a number of historic sites. I set up the calculations, by the way, in Microsoft Excel, with a template for 65 assemblages and 12 variables. I stopped there, frankly,
because I got bored out of my skull filling in all those cells to do the calculations. Anyone who’d like to play with these statistics is welcome to email me and I’ll attach the Excel template file.

I can go in a couple of directions here. For one thing, I can look within a site at those deposits that have similarity coefficients below one or two standard deviations of the average, and pick out anomalous test units to see what makes them special.

I can also construct similarity matrices for regional sites, which gets back to the problem of pattern and phase.

By default, I take South’s data for the sites that defined his Carolina pattern as a provisional standard for variability. For his five sites, the mean Brainerd-Robinson coefficient — the index of homogeneity — is 163.72 with a standard deviation of 10.60. That seems to be a high degree of similarity and a fairly tight variation. Provisionally, then, sites of comparable statistics might be said to belong to a pattern, but sites with a coefficient of similarity less than 153 could be said to be significantly more variable than the Carolina pattern. But — how does that compare to other sets of samples? I have eight sites in western Kentucky and Southeast Missouri that I can use as a comparative sample. The five western Kentucky sites were occupied after 1840, into the mid-
to-late 20th century. They had inhabitants with names like Smith, Moore, Sullivan, Morrill, and Tilghman, which I’ll take as a basic British-American heritage. The Southeast Missouri sites are in Ste. Genevieve: the Valle house had inhabitants of French descent, the Shaw house British-American, and the Delassus-Kern house German. With one exception the deposits belong to the same period as the western Kentucky sites (there’s a feature in the Valle sample that has a fill dating mainly to ca. 1790-1830). All the sites belong to middle- to upper middle-class families. With the possible perturbation of the French and German heritage, the sites should be fairly comparable. The data are recorded in the same functional assemblage groups as the Carolina pattern sites.

First, the five western Kentucky sites: they have an average BR coefficient of 157.64 ± 24.01. The average is within the Carolina index range, but the range of variability is wider.

For the three Ste. Genevieve sites, the index of homogeneity is 161.20 ± 20.92. Again the average is similar to that of the Carolina sites, and the range wider. The western Kentucky and Southeast Missouri sites form communities about as similar as, but somewhat more variable than, the sites that defined the Carolina pattern.
If I calculate the western Kentucky and Ste. Genevieve sites together, the index is $162.14 \pm 22.08$. The eight sites form as cohesive a community as each subgroup, which suggests that the French and German heritage has been submerged in a general American culture in this period, this region, and by the measure of this level of assemblage analysis.

Now, what happens when I compare the Carolina sites and my samples: I get an index of homogeneity of $136.94 \pm 34.60$. The mean is below the range of any of the regional groups, and the standard of deviation is half again as great. I was rather excited about this difference until I realized that there’s a big problem in making this comparison: comparability of data collection techniques. Field techniques are fairly standardized, and screen size is usually specified in reports. But lab procedures may not be quite as standard. I sort ceramics on a half-inch screen, and set aside the smaller sherds and glass. Many reports specify the number of sherds smaller than a threshold, which can be accounted for, but some do not. Also, many reports specify major type categories, but when describing the “other” sherds, do not attribute all of them to units or excavation levels, which can make statistics difficult. I also wonder if we count, say, nails differently: if I count all fragments that look like a nail, and someone else might
count only whole nails, or nails with heads, perhaps. I would say that only assemblages for which the lab procedures can be shown to be compatible can be compared in this way, and any calculation comparing projects for which the data collection and recording methods may not be the same must be taken with suspicion. There probably is a pattern distinction between eighteenth-century Carolina and nineteenth-to-twentieth century Mississippi Valley sites, but this index may not fairly characterize it. So, to say that one of my sites fits or does not fit the Carolina pattern may be neither accurate nor useful.

I don’t have time today to present data and indices for ceramics and glass assemblages, but exploring them with the same technique shows ceramic decorations to be less variable than the group patterning (except for a time dimension), but ware type and glass assemblages to be more variable. That’s interesting, but I don’t know yet where it’s leading me.

The point of all this, really, is that we can use many of the same techniques to explore assemblage variability in prehistoric and historic sites, in patterns and phases. I am not, emphatically not, suggesting that historical archaeologists should begin to define phases. The scale of the societies we’re dealing with is completely inappropriate to the phase concept, and we have other means and dimensions for assigning people to cultural groups.
But whether for phase or pattern, we need to explore assemblage variation in a number of ways, and avoid designating one type component and trying to create units based on unrigorous comparisons to that assemblage. Also, we need to build variability in different kinds of assemblages, not just functional group frequencies, into assemblage pattern studies.

The concepts of phase and pattern, in the sense that they emphasize and delve into assemblage similarities and dissimilarities, have points of congruence that allow for methods that are common to the subdisciplines. They are methods of archaeology, not historical or prehistoric archaeology.

Thank you.