CHAPTER SIX

Alternative Analyses

By C. Roger Nance, Jan de Leeuw, Kathleen Prado, and David S. Verity

Introduction

In the previous chapter, we detailed type distributions through a proposed sequence based on a single analysis. At this point, we should address the following question: how would a different analyses or the use of different variables (types) affect the outcome? To what extent would different progressions of types or sherd lots materialize? We explored this issue through four additional analyses, three of them employing CA and the last the EDM method. Both of these statistical procedures are described in chapter 2.

Correspondence Analysis of Twenty-three Types: CA23

For this analysis, twenty-three types (listed in table 6-1) were selected based on their high frequencies, with type frequencies in the study sample varying between 78 and 1,270 sherds. A few types, those residual in nature, such as Gray/Buff Utility (1) and Eroded Fine Ware (4), were excluded. Figure 6-1 portrays the two-dimensional distribution of types, and inspection reveals a close parallel between this CA and the original three-site model (see fig. 4-7). In the original CA, Huistla Polychrome sherds were combined into two groups and occurred at the late end of the sequence. In figure 6-1, included Huistla Polychrome types (7, 8, 19, 20, 21) likewise cluster in the same relative position (although here the

time line is running in the opposite direction). Likewise, at the early end of both sequences occur Red on Buff, Early, and its equivalent, Red on Buff, Complex (24); three White on Red types (16, 17, 18) that were scattered in early positions in CA23 but which combined together in the original three-site CA; and type 15, Incised Polychrome. Near the center of both grids are Red on Buff, Late, combined in the original three-site analysis but occurring as types 22 and 23 for CA23; the Comales type (2); and Gray/Buff Slipped and Polished (5). The type Black/Brown Slipped and Polished (6) can be found in slightly later positions on both grids.

Table 6-1: Types in the additional analyses

Type Code	CA23	CA17	CA40	EDM23	Frequency	Name
2	+	+	+	+	1270	Comales
3	+	+	+	+	1254	Grater Bowls, General Form
5	+	+	+	+	314	Gray/Buff Slipped and Polished
6	+	+	+	+	325	Black/Brown Slipped and Polished
7	+		+	+	86	Grater Bowls, Huistla Polychrome, Complex
8	+		+	+	79	Grater Bowls, Huistla Polychrome, Parallel Lines
10			+		20	Grater Bowls, Black Band at Rim
11			+		32	Grater Bowls, Red on Buff
13	+		+	+	83	Grater Bowls, Fine Ware

14	+	+	+	+	249	Black and White on Red
15	+	+	+	+	269	Incised Polychrome
16	+	+	+	+	324	White on Red, Simple
17	+	+	+	+	157	White on Red, Complex
18	+		+	+	79	White on Red, Broad Stripe Outlined
19	+	+	+	+	148	Huistla Polychrome, Complex (zigzag)
20	+		+	+	78	Huistla Polychrome, Complex (unique)
21	+	+	+	+	245	Huistla Polychrome, Stripes
22	+	+	+	+	287	Red on Cream, Thin Parallel Lines
23	+	+	+	+	365	Red on Cream, Broad Parallel Lines
24	+	+	+	+	264	Red on Cream, Complex Design
26	+		+	+	92	Polychrome, White Dots
40			+		58	Fine Engraved, Arcaded
41			+		46	Historic, Glazed Majolica
42			+		38	Washboard Exterior
43			+		64	Brushed or Combed Utility

44			+		43	Hatched from Rim, Incised
45			+		37	Black Crude Engraved, Complex
46			+		14	Black Crude Engraved, Suspended Triangles
47			+		41	Dark Red to Black Engraved, Complex
48			+		25	Reed Punctate
49			+		21	Exterior Incised, Deep Parallel Lines
50			+		13	Red Fine Incised
52			+		49	Miscellaneous Broad-Line Incised
53			+		50	Complex Broad-Line Incised
61	+	+	+	+	152	Fine Red Ware
62	+	+	+	+	680	Red Utility Ware
63	+	+	+	+	360	Red-Painted Jars, Streaky Polish
64	+	+	+	+	208	Red Utility, Very Thick
90			+		99	Black and Red Types (33, 35, 37, 38)

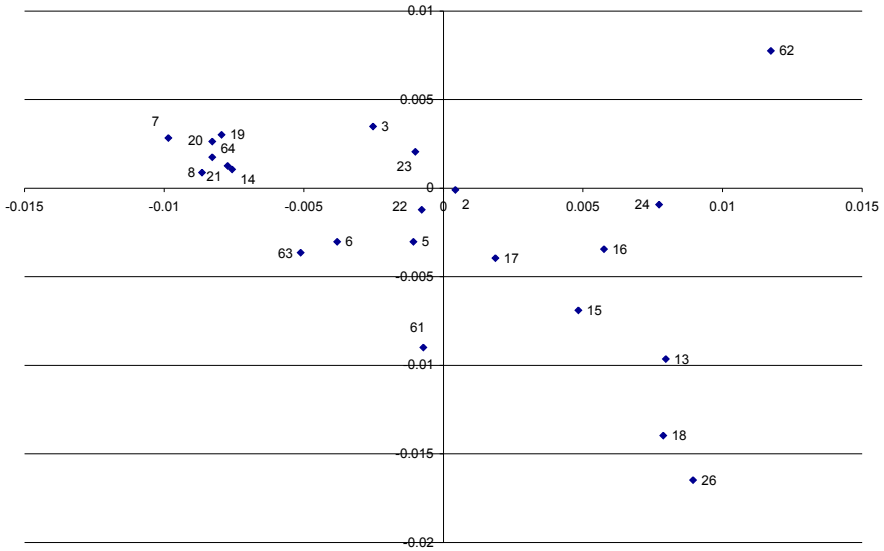

Figure 6-1. Type plot, CA23.

One interesting feature of figure 6-1 is the close proximity of points near the late end of the sequence and the fact that they become more dispersed the earlier their chronological positions. That is, in this analysis early types have more disparate distributions relative to one another than are found among their later counterparts. The same phenomenon can be seen operating in figure 6-2, which shows the plot for lots or samples of sherds in this same twenty-three-type analysis. The reason for this result is unknown, but it could be that early pottery on the site represents culturally distinct groups residing on the site at different periods of time and not continuous occupation by a single culture, as might have occurred later. Figure 6-2 identifies each lot by site, and, as in the three-site lot graph (see fig. 4-10), Tiana loci are all in early positions, in this case, all having positive x-axis values.

Another way to assess the degree of difference between the original CA and CA23 is to compare type median CAvalues from original CA lots, as presented in chapter 5, to type x-axis values of CA23, on a type-by-type basis. This comparison is portrayed graphically in figure 6-3, and one can see a strong negative correlation—negative, again, because

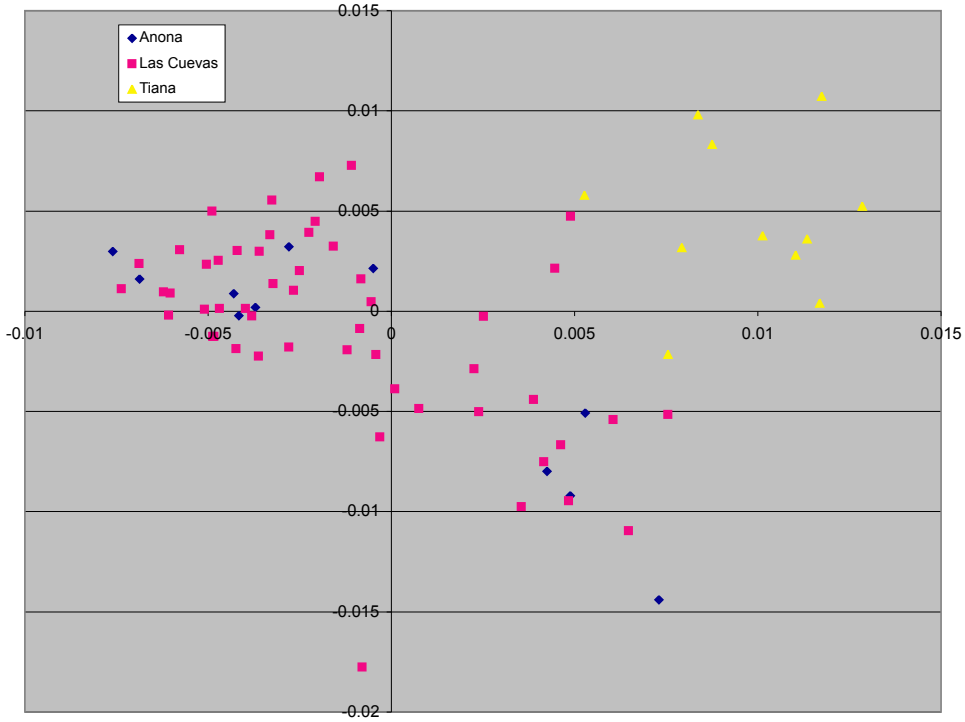

Figure 6-2. Sample plot, CA23.

the two CAs are aligned in opposite directions. The Pearson correlation coefficient of r = -.9383 has a significance level of ***; the two type progressions are strongly correlated.

The question remains, how well would CA23 serve in revealing the distribution of types in terms of its own lot (x-axis) CAvalues and presumably through time? Figure 6-4 shows these distributions for the six grater bowl types, although here the signs were reversed for CAvalues, which reverses the direction of the type distributions. This figure, then, is comparable to the similar grater bowl summary graph included for x-axis values from the original three-site CA (see fig. 5-1). For the CA23 graph, type distributions do have approximately the same relative positions, but separations are not as clear and distributions are not as consistent, that is, we see more outliers. Part of the problem is the smaller scale, -1.5 to 1.0 in figure 6-4 compared to -2.5 to 2.0 in figure 5-1.

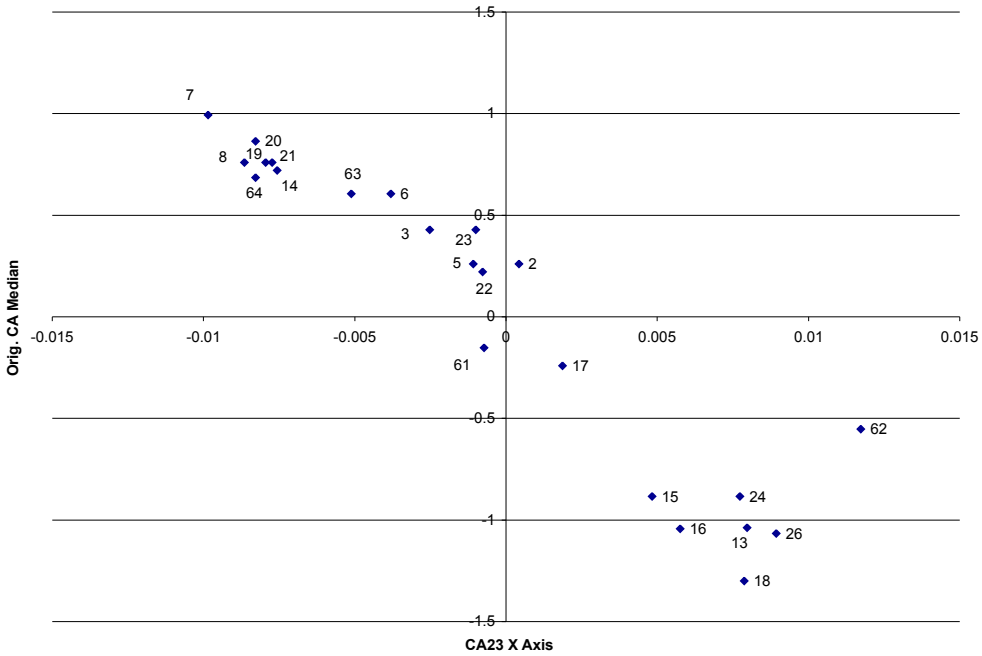

Figure 6-3. CA23 x-axis type value by type median from original CA lot values.

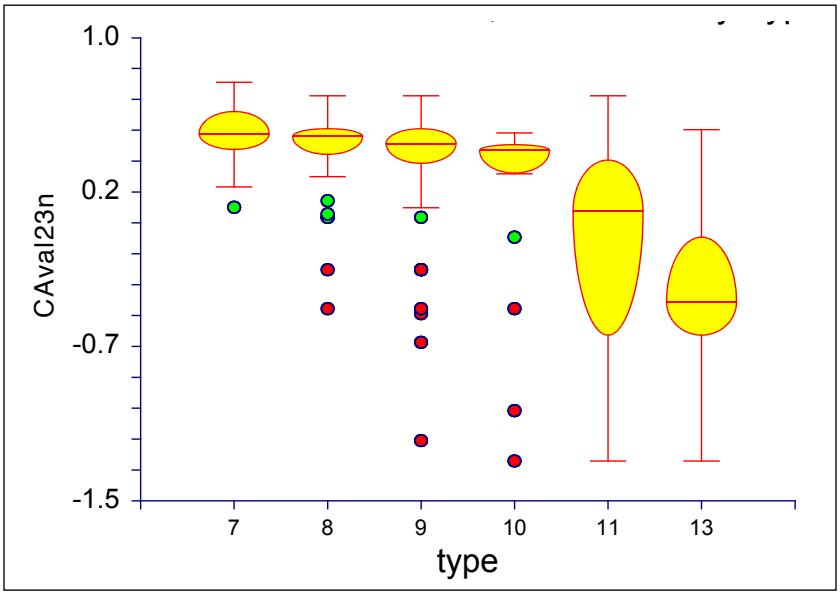

Figure 6-4. CA23 box plots, grater bowl lot CAvalue distributions by type.

In sum, CA23 does show very much the same sequence found in the original CA, but the former does not appear as useful in demonstrating the full range of type distributions. From the archaeologist's perspective, this lesser utility could be due to the inclusions of the red ware types 61–64. While high in frequency, sherds of these types proved difficult to classify, and these types might have had a destabilizing influence. One indication of this effect is that the type most out of alignment in figure 6-3 is type 62, Red Utility Ware.

Correspondence Analysis of Seventeen Types: CA17

To reduce the problem of sampling error, we also considered another CA in which included types were limited to those with at least 148 sherds. All seventeen of these types (listed in table 6-1) also figured in the twenty-three-type CA discussed previously.

Figure 6-5 portrays the two-dimensional grid for CA17 types. A comparison with the original three-site CA (see fig. 4-7) shows them to

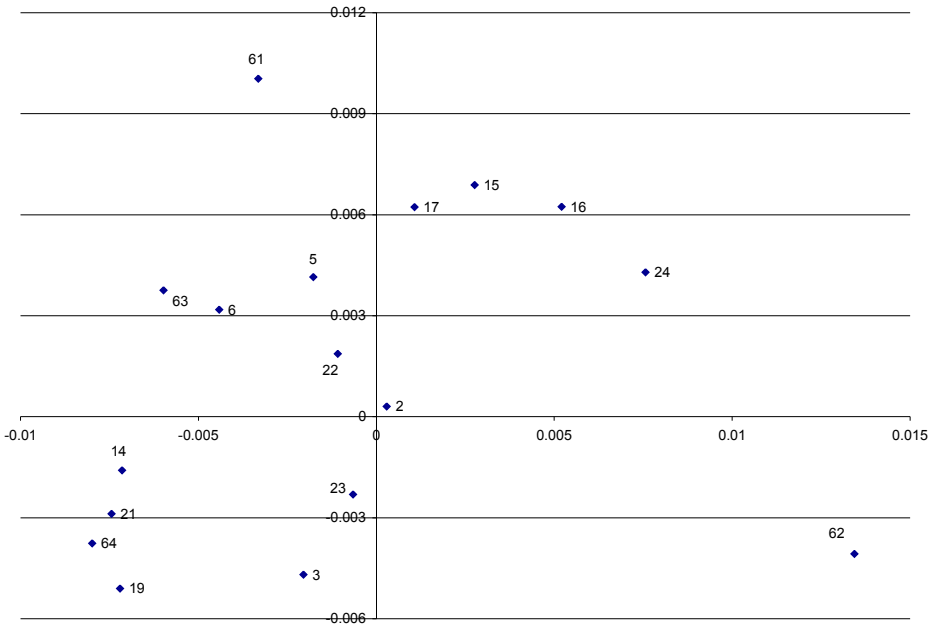

Figure 6-5. Type plot, CA17.

be close, albeit the two type progressions again run in opposite directions. In figure 6-5, two Huistla Polychrome types (14 and 21) anchor the late end of the sequence; Gray/Buff Slipped and Polished (5), Comales (2), and the two late Red on Buff types (22 and 23) occupy intermediate positions; and two White on Red types (16 and 17), Incised Polychrome (15), and the early Red on Buff type (24) are early. The type Black/Brown Slipped and Polished (6), is situated between the Huistla Polychrome and intermediate types. One can see essentially the same progression in the two figures.

Figure 6-6 contains the two-dimensional plot for lots in CA17. Here we see the same increasing density of points from early to late found in CA23 (see fig. 6-2), and to some extent the same holds for type plots (cf., figs. 6-5 and 6-1). Another comparison involves the median type values for the seventeen types taken from the original CA, in which lot x-axis values were applied to all sherds by type, compared to type x-axis values from CA17 (fig. 6-7). Again, we see a strong correlation (r = -.8431, significance level = ***) between the two type alignments. A

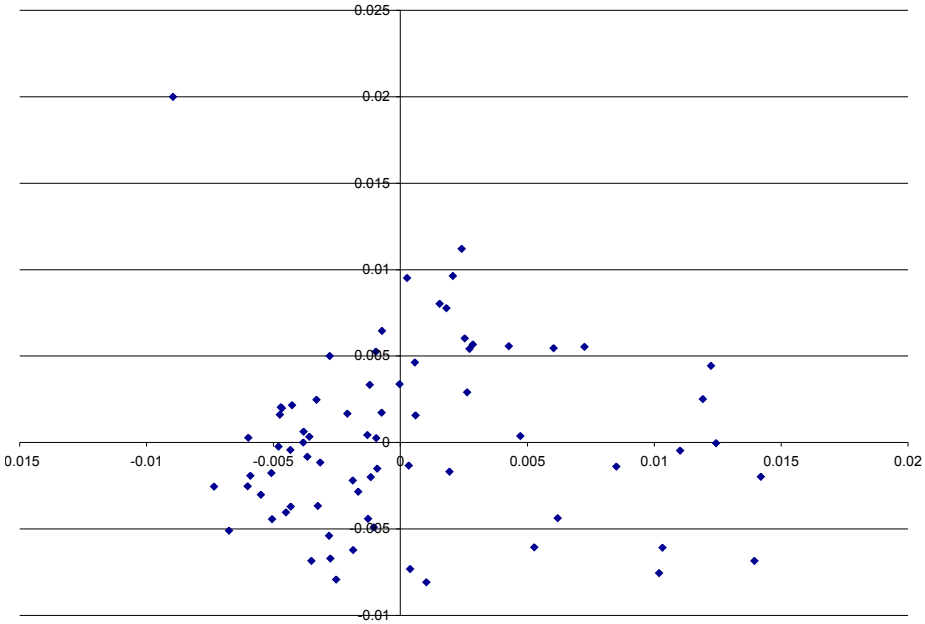

Figure 6-6. Sample plot, CA17.

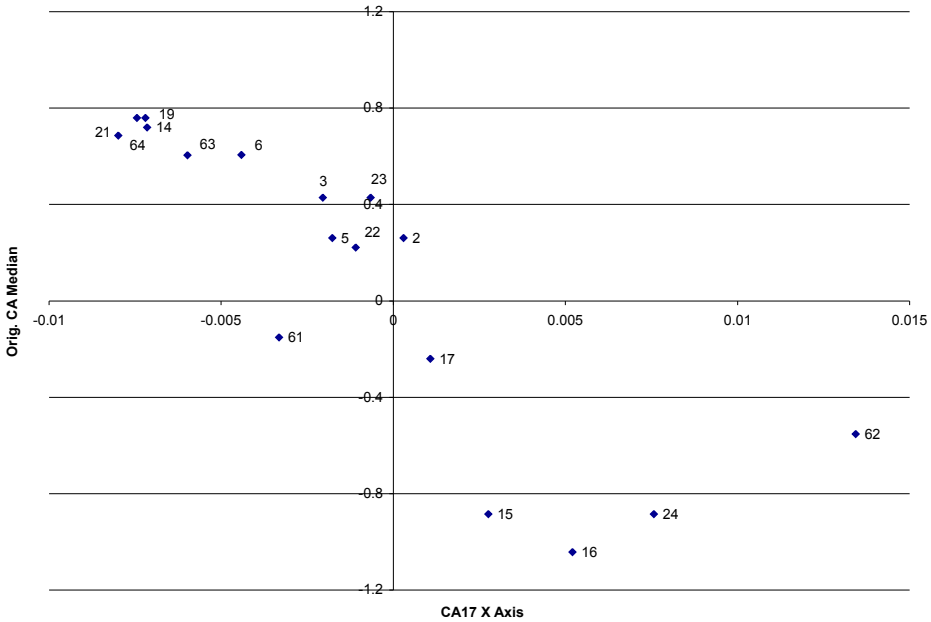

Figure 6-7. CA17 x-axis type value by type median from original CA lot values.

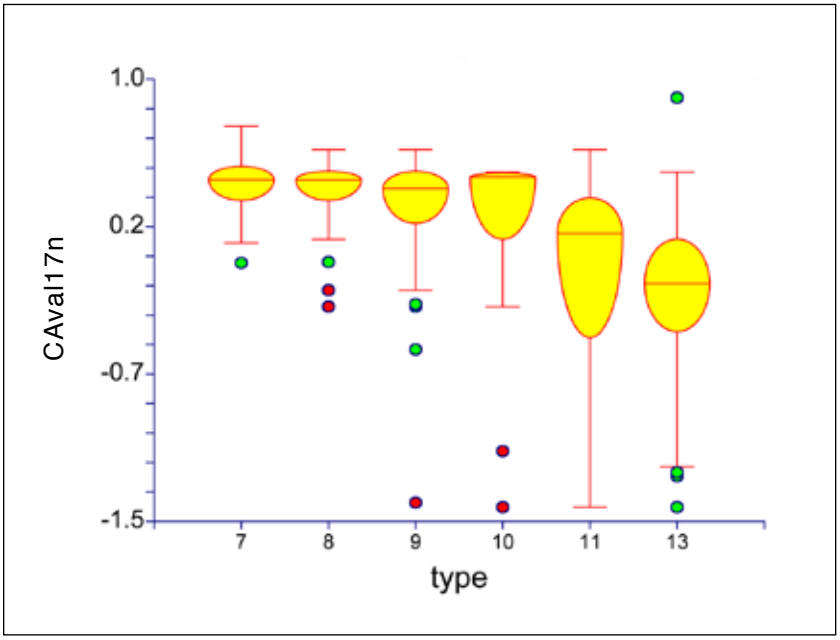

Figure 6-8. CA17 box plots, grater bowl lot CAvalue distributions by type.

final comparative view involves lot x-axis values for CA17, which were summarized for sherds by type. Again, grater bowl types serve as an example (fig. 6-8) to let us evaluate how well this approach might depict a detailed picture of type distributions. Compared to figure 5-1, this rendition of grater bowl type distributions appears less satisfactory, similar to our finding for CA23.

Correspondence Analysis of Forty Types: CA40

In this third alternative analysis, we maximized the number of types included. Excluding residual types, we included all those with at least thirteen sherds, if they showed some tendency to cluster timewise in the original CA sequence. The forty included types are listed in table 6-1. One composite type is included, type 90, which contains data for four types (33, 35, 37, and 38). Because of the inclusion of types with low frequencies, we anticipated some volatility, and this volatility was revealed in the original two-dimensional type plot (fig. 6-9). Here, most type loci tend to be aligned roughly parallel to the x-axis, but five types can be identified as outliers: 46, 49, 52, 53, and 61. Most of these outlier

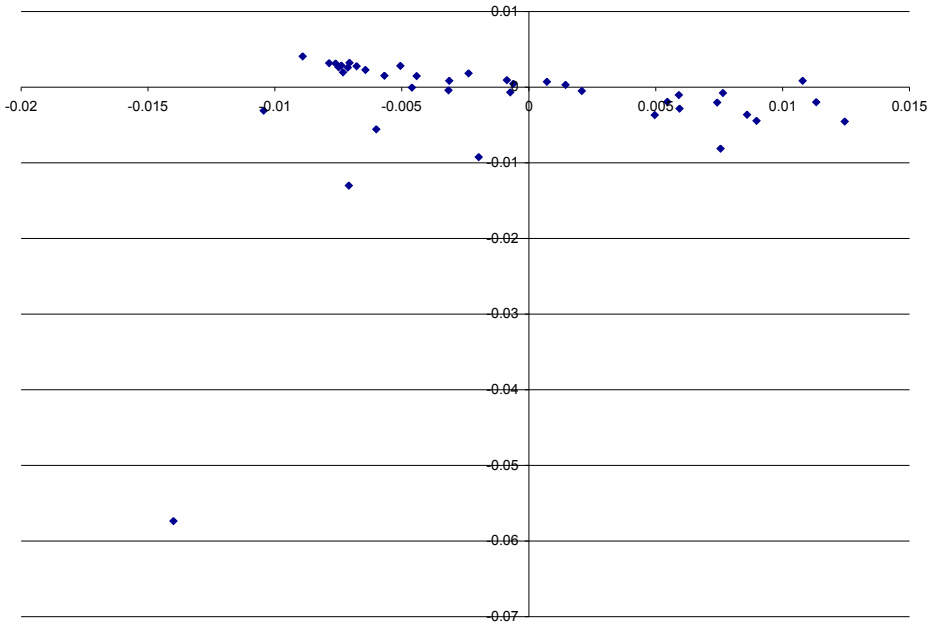

Figure 6-9. Type plot, CA40 (with outliers).

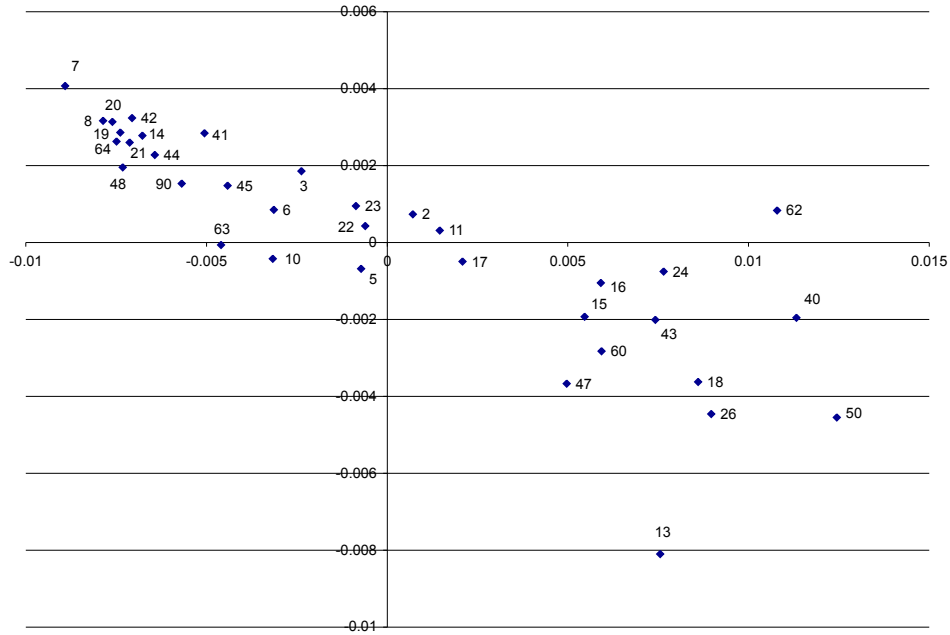

Figure 6-10. Type plot, CA40 (outliers removed).

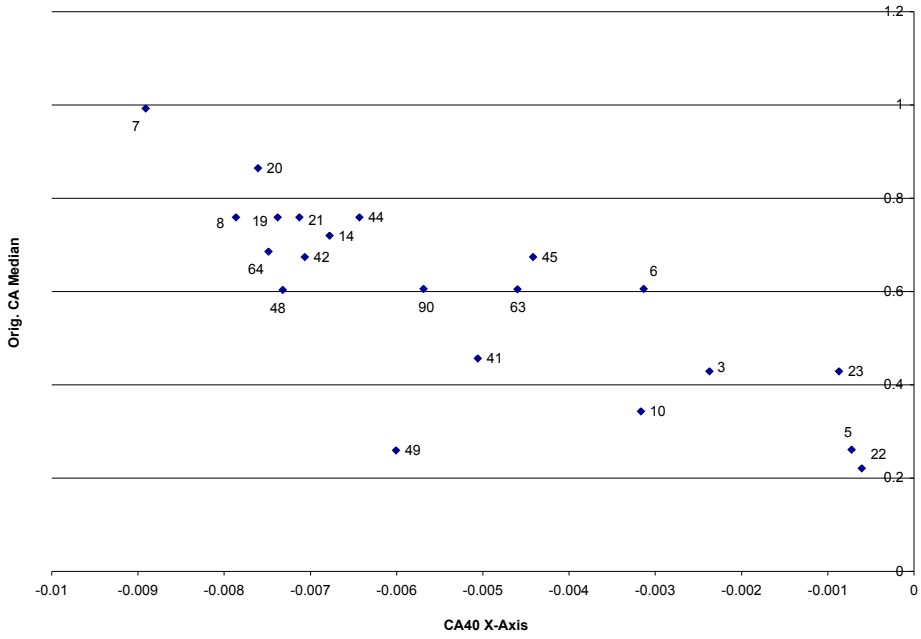

Figure 6-11. CA40 x-axis type value by type median from original CA lot values (late types).

types have low frequencies (fourteen to fifty sherds each in the data set). The exception, type 61, is a red ware type, and its reliability was suspect at the outset. We removed these types from further consideration and focused our attention on the remaining thirty-five types.

Figure 6-10, then, depicts the same plot but with the five outliers removed. Again, we see compatibility with the original three-site CA type plot (see fig. 4-7), and types line up consistently, as reported for CA23 and CA17; the type progression once more moves in the direction opposite that of the original. In figure 6-11, medians for type distributions, which we generated from the original CA lot x-axis values, are compared to type x-axis values for CA40 in a scattergram. Figure 6-11 depicts only late types, but when data for all forty points were compared, they were found to be highly correlated (r = -.972, significance level = ***).

It also is instructive to consider the lot distributions in the two-dimensional grid for CA40. As for types, we found the overall picture

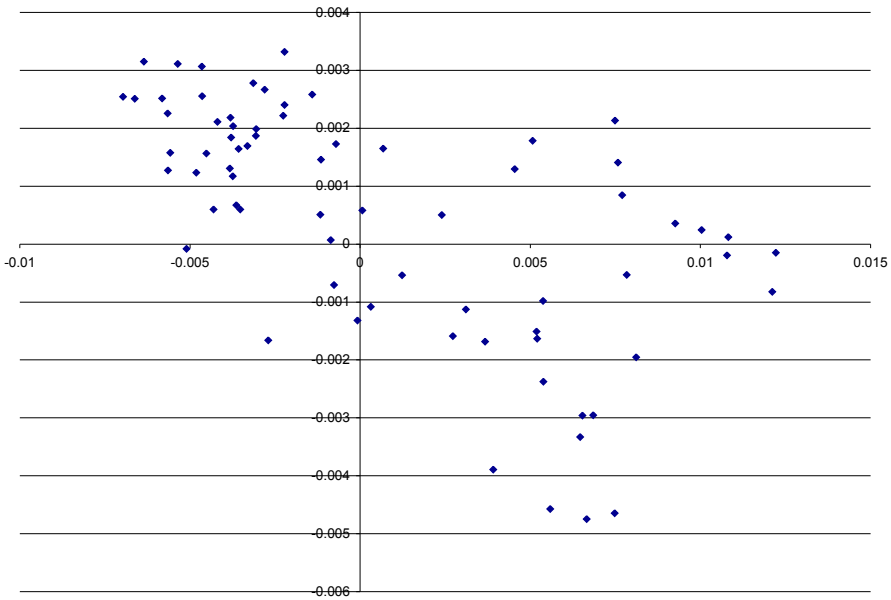

Figure 6-12. Sample plots, CA40 (outliers removed).

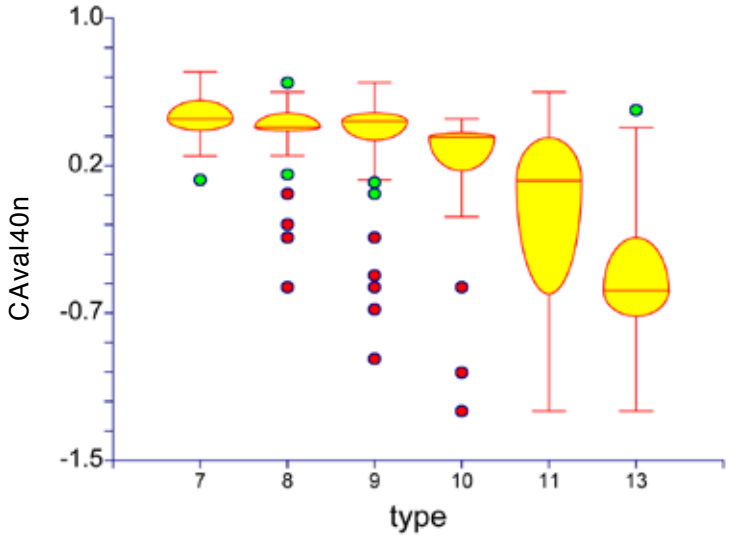

Figure 6-13. CA40 box plots, grater bowl lot CAvalue distributions by type.

somewhat distorted by the presence of outlier lots, but after the removal of three obvious outliers, the pattern of lots for CA40 (fig. 6-12) was similar to those for CA23 and CA17: closely spaced loci on the later portion of the grid and dispersed loci on the earlier portion.

Finally, we generated another box plot for the six grater bowl types, here using x-axis lot values from CA40 to summarize type distributions (fig. 6-13). As for CA23 and CA17, the results still seem inferior to those discovered in the original CA (see fig. 5-1).

Exponential Distance Model Analysis: EDM23

In this last alternative analysis, the EDM procedure was applied to the same twenty-three types included in CA23 (see table 6-1). This method produced a much more even array of types in the two-dimensional grid (fig. 6-14) without an increase in loci clustering through time. Also in this case, the type progression runs in the same direction as the original three-site CA: early types with negative x-axis values and late types

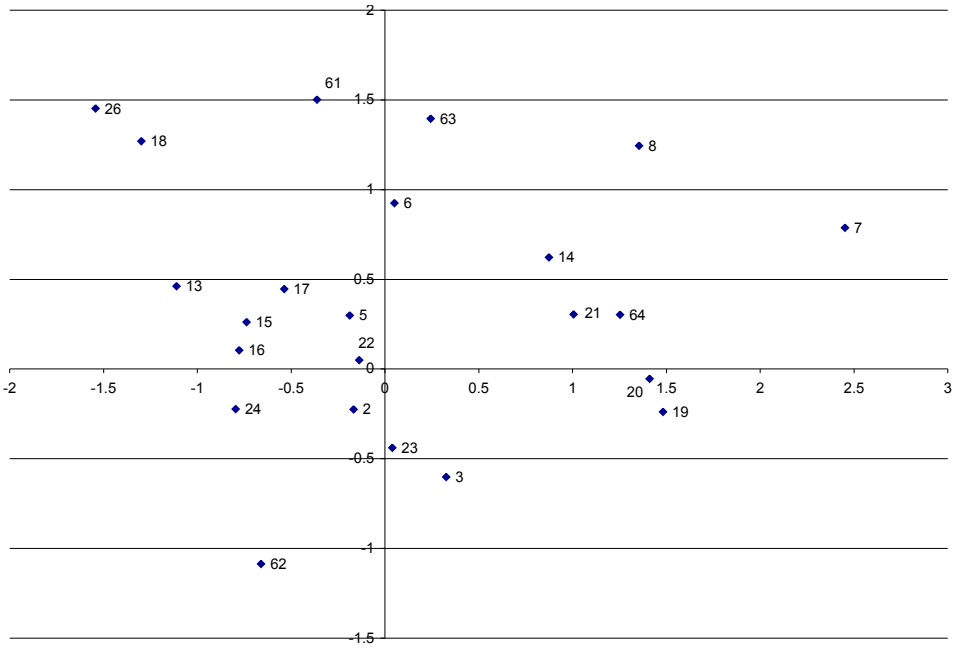

Figure 6-14. Type plot, EDM23.

with positive. Regarding the content of the two type progressions, parallels are consistent with those identified for the three previously described alternative CAs.

Figure 6-15 displays lot distributions for EDM23. As for types, we see no obvious evidence of an increased cluster of loci through the sequence. In this graph, loci are identified by site, and the early position of Tiana lots in the grid is consistent with findings from the original three-site CA (see fig. 4-10) as well as CA23 (see fig. 6-2).

We compared EDM23 type x-axis distributions to median values from the original CA lot summaries by type (fig. 6-16) and type x-axis distributions for CA23 (fig. 6-17). Correlations for both are high, negative in the case of CA23 and EDM23, where type orders are reversed. Relationships between these two CAs and the EDM analyses are curvilinear, and the order of types, one compared to the other, is remarkably similar. The EDM23 box plot (fig. 6-18) for grater bowl types is similar

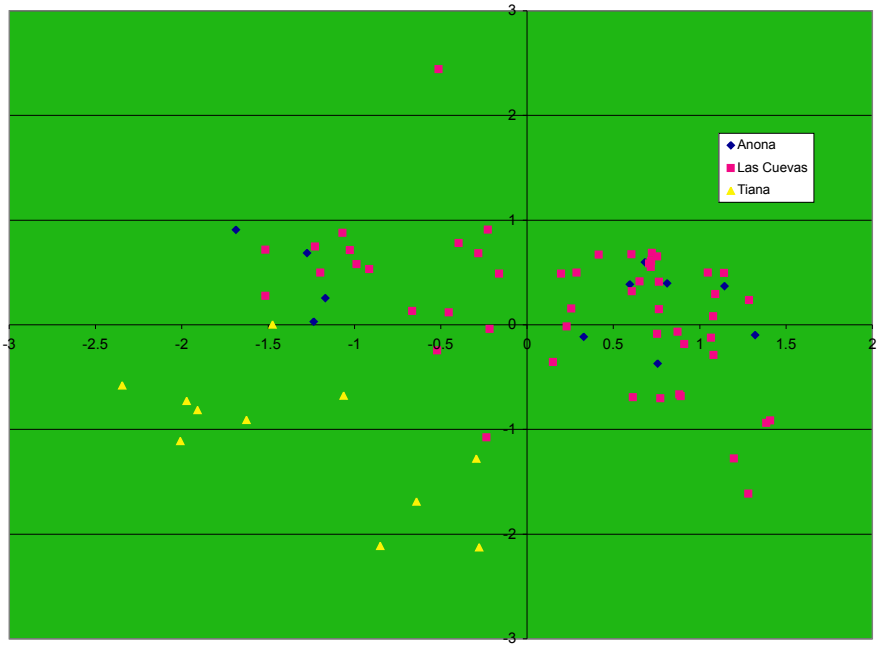

Figure 6-15. Sample plot, EDM23.

to but less satisfactory than that for the original CA; it displays more outliers and less spacing of types.

Summary

The aim of this brief chapter has been to explore the possibility that the original three-site CA and the site-specific CAs discussed earlier were somehow influenced by the selection of types for the analysis. In attempting to avoid this potential source of bias, we generated two additional CAs with type selection based only on type frequency: all types with more than 148 sherds (CA17) and all types with more than 78 sherds (CA23). In a third CA, we circumvented the selection process by including as many types as possible (CA40). The fourth analysis (EDM23) employed a different statistical procedure. Yet in all four of these alternative analyses, we found evidence for generally the same alignment of types. We can conclude, at least based on available findings, that this type alignment is embedded in the data analyzed.

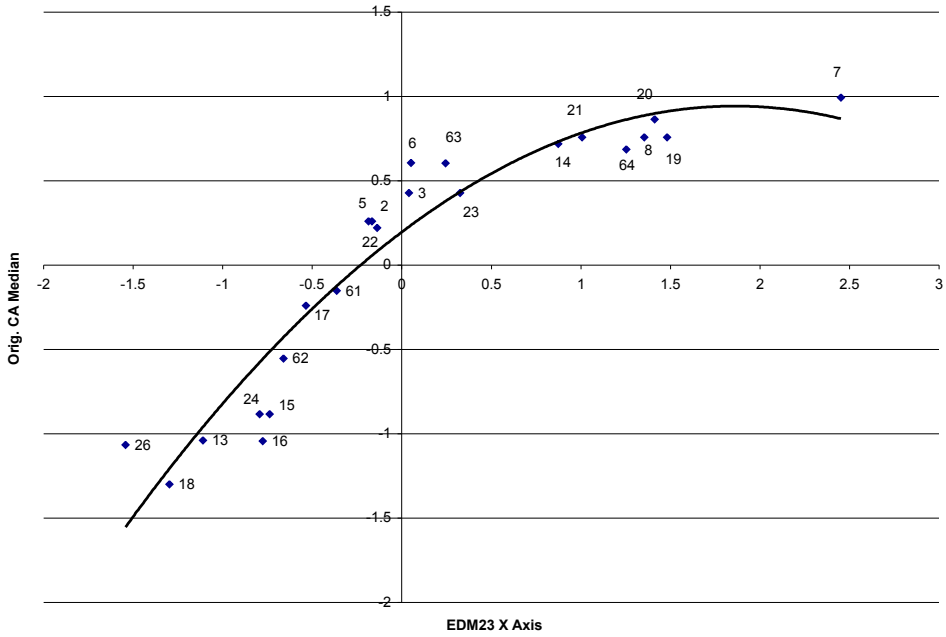

Figure 6-16. EDM23 x-axis type value by type median from original CA lot values.

A second result of this exercise was the discovery in analyses CA23, CA17, and CA40 of the clustering of lots near the later end of the sequence and their greater dispersal in earlier samples, showing more diverse distributions of early types. This finding may relate to an increase in settlement size or duration through time at these sites, but much more fieldwork and analysis will be required to explore this issue.

Finally, the four alternative analyses do say something about the actual distributions of types through the sequence, not just their relative positions on a CA grid. By applying lot CAvalues to all sherds in the study, not just those figuring in the CA, we could at least suggest the nature of the distribution for most individual types. Parallels across the different analyses can be seen in the four box plots included in this chapter.

Another way to glimpse similarities and differences in type distributions among the analyses has to do with skewness indices in type distributions. In chapter 5, we showed a positive correlation in the original three-site CA between a type's median CAvalue and the skewness value

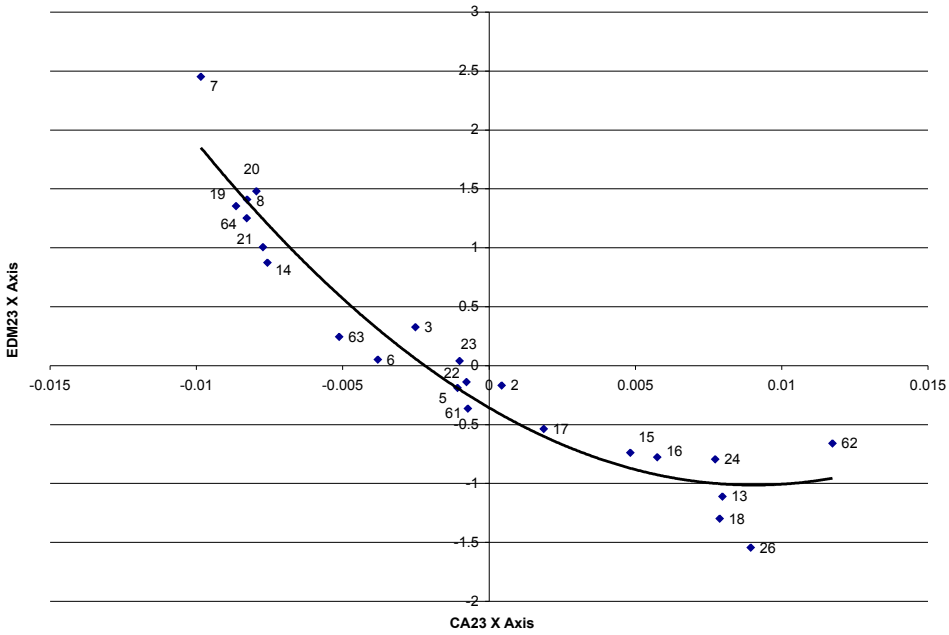

Figure 6-17. EDM23 x-axis type value by CA23 x-axis type value.

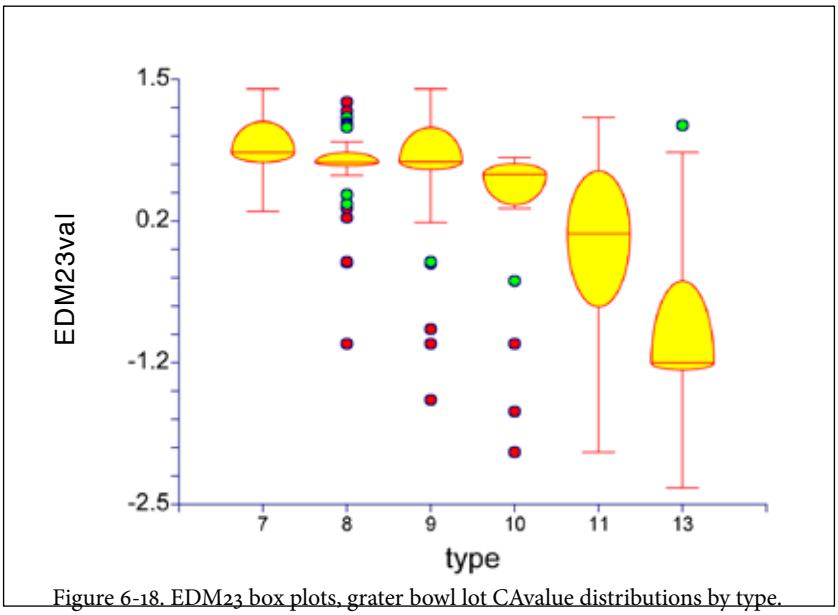

Figure 6-18. EDM23 box plots, grater bowl lot CAvalue distributions by type.

of its distribution pattern. Early types tend to have positive skewness values and late types, negative values. Without delving too far into the intricacies of these alternative analyses, we included skewness indices for six of the incised/engraved types for each of the five analyses under consideration here (table 6-2). Keeping in mind that three of the type progressions are running in opposite directions to the original CA and that skewness value signs should be reversed in those cases, we can see evidence of correspondence for three of the four alternative analyses.

Table 6-2: Type skewness by alternative analysis

Age	Type	Original CA	CA17	CA23	CA40	EDM23
Late	44	-.518	2.677	2.416	2.379	-1.311
Late	45	-.566	1.387	1.237	1.343	-1.256
Late	46	-.879	.790	.865	.895	-1.004
Early	40	.811	.042	-.403	-.571	.574
Early	43	.470	.962	-.906	-.551	.351
Early	51	.593	.147	-.133	-.112	.122