4

Multivariate statistics and assemblage comparison

Roberta Tomber
Department of Archaeology, University of Southampton

4.1 Introduction

In Britain, the collection of quantified ceramic data is standard practise, so much so that the different methods of quantification have been rigorously tested and assessed (Millett 1979, Orton 1975, Orton 1982). The most reliable measure overall is that of estimated vessel equivalents (eves) (Orton 1975) and some recent publications rely on this measurement alone (Keeley 1986). In contrast, in the Mediterranean quantification of archaeological data is only just gaining popularity and is still treated with scepticism by many. As for there being any sort of consensus on the type of quantification used, a perusal of the literature indicates that four recent publications used that many different methods of quantification (Keay 1984, Fulford & Peacock 1984, Riley 1981, Whitehouse *et al.* 1985). However, it is unfair to level these criticisms without acknowledging the background to this situation. The quantities recovered on a Roman or Byzantine site are vast, with a six week season easily producing three or four metric tonnes of material. Because of the prohibitive cost and time necessary to process fully such amounts of pottery, the tradition has been to take an art historical approach and avoid a comprehensive study, instead selecting particular classes of material.

Now that Mediterranean archaeologists are beginning to collect quantified data (normally counts recorded as some permutation of rims, bases, handles or body sherds, or less frequently weight) it is essential to find a valid way in which to utilize this information. Assemblage comparison is usually made between percentages of pottery types within a deposit and these then form the basis for economic inference. However, the lack of understanding of intra-deposit variability means that it is difficult to separate 'significant' differences from those which can be attributed to an acceptable range of within deposit variability.

This paper attempts to test the degree of homogeneity and define the main sources of variability within a given assemblage through Principal Components Analysis. A method of data presentation and comparison, using Discriminant Analysis, which takes these aspects of variability into account and assesses similarity between deposits will then be suggested.

Statistical analysis was carried out on the University of Southampton's mainframe computer. Most of the work was completed on an ICL 2976; but since January has been on an IBM 3090. Both Principal Components and Discriminant Analysis were done with SPSSX Version 2 (SPSS 1983); and plotting used the Clustan Scatter procedure (Wishart 1978).

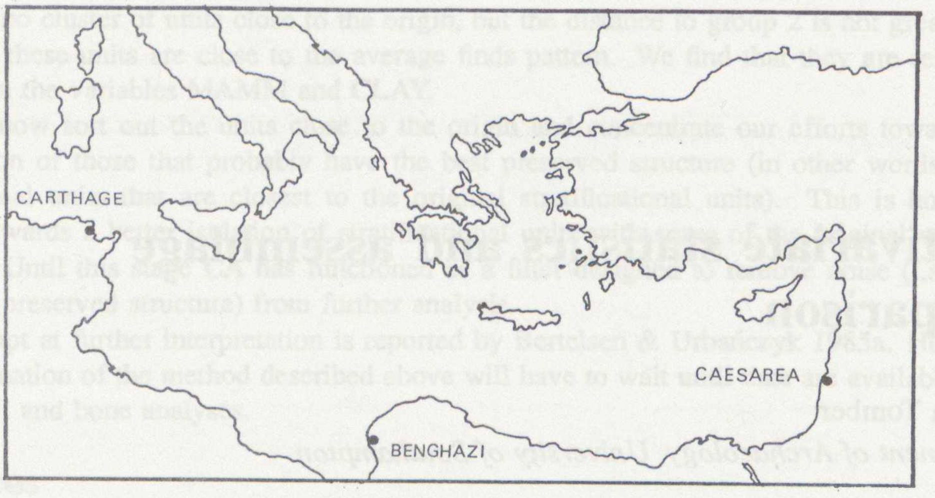

Fig. 4.1: Location map of Carthage and other sites included in the analysis

4.2 The case study

The data used for this investigation from the 1983 excavations at the Carthage circus (Fig. 4.1) and one deposit was considered especially suitable for examining variability. The material is primarily late-4th to early-5th century AD, although coin evidence places its deposition in the late-5th century AD. It provided an ideal test group, having been dug stratigraphically and producing 88 separate ceramic layers which could be compared with each other. Furthermore, the pottery from individual stratigraphic contexts varied from c. 40 to 150,000 grammes, allowing questions of sample size to be addressed. During sorting and classification the layers appeared generally homogeneous, but when percentages of individual types were calculated a diverse range of values emerged and the best method of assessing these differences was not clear.

4.3 Principal Components Analysis

One of the ultimate aims of the study was to define sources of variability within a single deposit. Therefore it was first necessary to test the assumption of homogeneity and this was done using Principal Components Analysis (see Doran & Hodson 1975, Shennan 1987) on weight values. Analysis included all 88 contexts and 104 ceramic types, *i.e.* all those which were represented by at least 100 grammes. The resulting matrix indicated a high degree of correlation between all contexts, and this was supported by the component loadings of which only four were considered to be poorly loaded, *i.e.* having values less than 0.67.

To try to define the sources of the variability that did exist, the assemblage was broken down into smaller groups based on functional ware types. These were defined as cooking wares, fine wares, undecorated table wares and amphorae. PCA was then run twice for each ware group, using both contexts and types as variables, so that a correlation matrix was created for each. As in the previous analysis this was done using weight data. The same general pattern could be seen for each ware group, and this is most clearly illustrated by the plot of component scores for plain table wares (Fig. 4.2).

As can be seen, most of the types and stratigraphic contexts cluster together, generally with

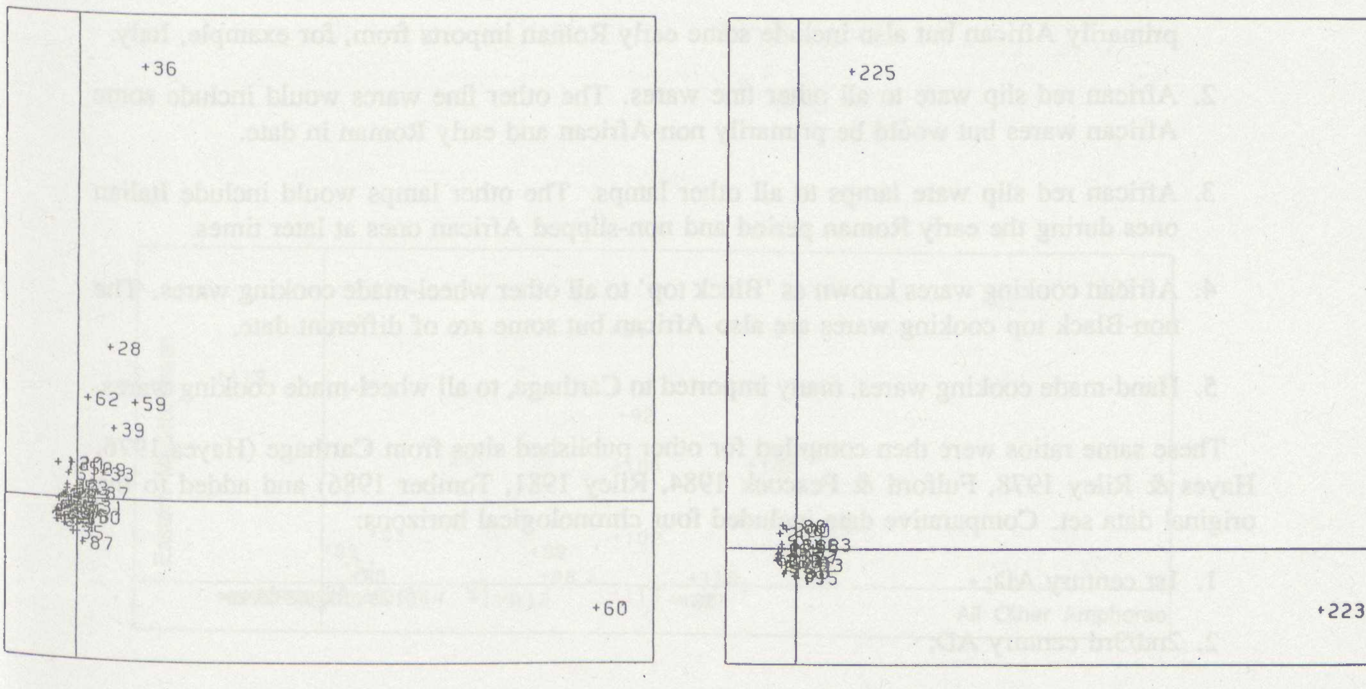

Fig. 4.2: Component scores for plain table wares, with contetxs and types as case identifiers

two extreme outliers. The two main outliers when vessel type is the case identifier are, without exception, the two most common forms for the particular ware group. If the two most common types are removed from the analysis the next two largest then become the outliers. The outlying stratigraphic contexts always include the largest one—Layer 60—which is four times greater than the next largest one.

By comparing the outlying types with the outlying layers two main sources of variability are revealed—assemblage size and ratios of types. In this case the ratio, by weight, between Type 223 and Type 225 in Layers 36 and 60 is very different, one being 1 to 5, the other 1 to 9. Layers 60 and 36 are both large contexts but are dispersed on the plot because of their different ratios. Component 1 is on the horizontal axis and relates to absolute sample size, while Component 2 is on the vertical axis and is the ratio between particular types.

In this way PCA provided a method for identifying assemblage size and ratio of types as two possible sources of variability. These were subsequently used as a basis for the next stage of analysis, assemblage comparison.

4.4 Assemblage comparison

Five ratios were calculated from weight to reflect chronological and long-distance economic trends. They were selected without reference to results from PCA but to incorporate what were known to be significant features of pottery assemblages for this period. The ratios are as follows:

1. Eastern Mediterranean amphorae (those types which are known in Britain as 'B wares' (Thomas 1959, Thomas 1981) to all other amphorae. These other amphorae would be

primarily African but also include some early Roman imports from, for example, Italy.

2. African red slip ware to all other fine wares. The other fine wares would include some African wares but would be primarily non-African and early Roman in date.

3. African red slip ware lamps to all other lamps. The other lamps would include Italian ones during the early Roman period and non-slipped African ones at later times.

4. African cooking wares known as 'Black top' to all other wheel-made cooking wares. The non-Black top cooking wares are also African but some are of different date.

5. Hand-made cooking wares, many imported to Carthage, to all wheel-made cooking wares.

These same ratios were then compiled for other published sites from Carthage (Hayes 1976, Hayes & Riley 1978, Fulford & Peacock 1984, Riley 1981, Tomber 1986) and added to the original data set. Comparative data included four chronological horizons:

1. 1st century AD;

2. 2nd/3rd century AD;

3. late Roman/early Vandal and

4. Byzantine.

The values for these ratios were then plotted. Fig. 4.3 shows eastern Mediterranean amphorae—the 'B' wares—on the vertical axis and all other amphorae on the horizontal, with each number referring to a separate deposit. Some basic trends are immediately visible. Sample size is important with the larger layers, such as 96, outlying. The bottom graph included all deposits; the top one excluded the largest ones since their size distorted the scale of the plot. Without the outliers the pattern emerges more clearly and therefore it is best to concentrate on the upper graph. Here cases 89–107 and 118–119 are all Byzantine in date, and most of these separate out from cases 1–88, the late Roman ones. While some of this distinction is based on assemblage size and there is a good deal of variability amongst them, the Byzantine deposits nevertheless display greater values on the vertical axis for eastern Mediterranean amphorae than all except one of the late Roman samples. Numbers 107–113 and 120 represent the 1st to 3rd century deposits, some of which separate out from the later ones. Plotting of ratios is an immediate way to illustrate gross differences between deposits but is limited to a single ratio and does not provide a rigorous method for comparing assemblages.

4.5 Discriminant Analysis

Discriminant Analysis (see Doran & Hodson 1975, Shennan 1987) was the best technique for approaching this problem. In this particular analysis the groups relate to chronology with each case representing an assemblage. Membership is known in every case. The SPSSX Stepwise selection technique was used with Mahalanobis distance.

Initial analysis was run using all samples from each of the four chronological groups. It was apparent from the start that there was substantial overlap between 1st and 5th century material. The value of many of the ratios for these two periods are identical although they hide more complex patterns. For example, both could frequently have a '0' value for the ratio of African

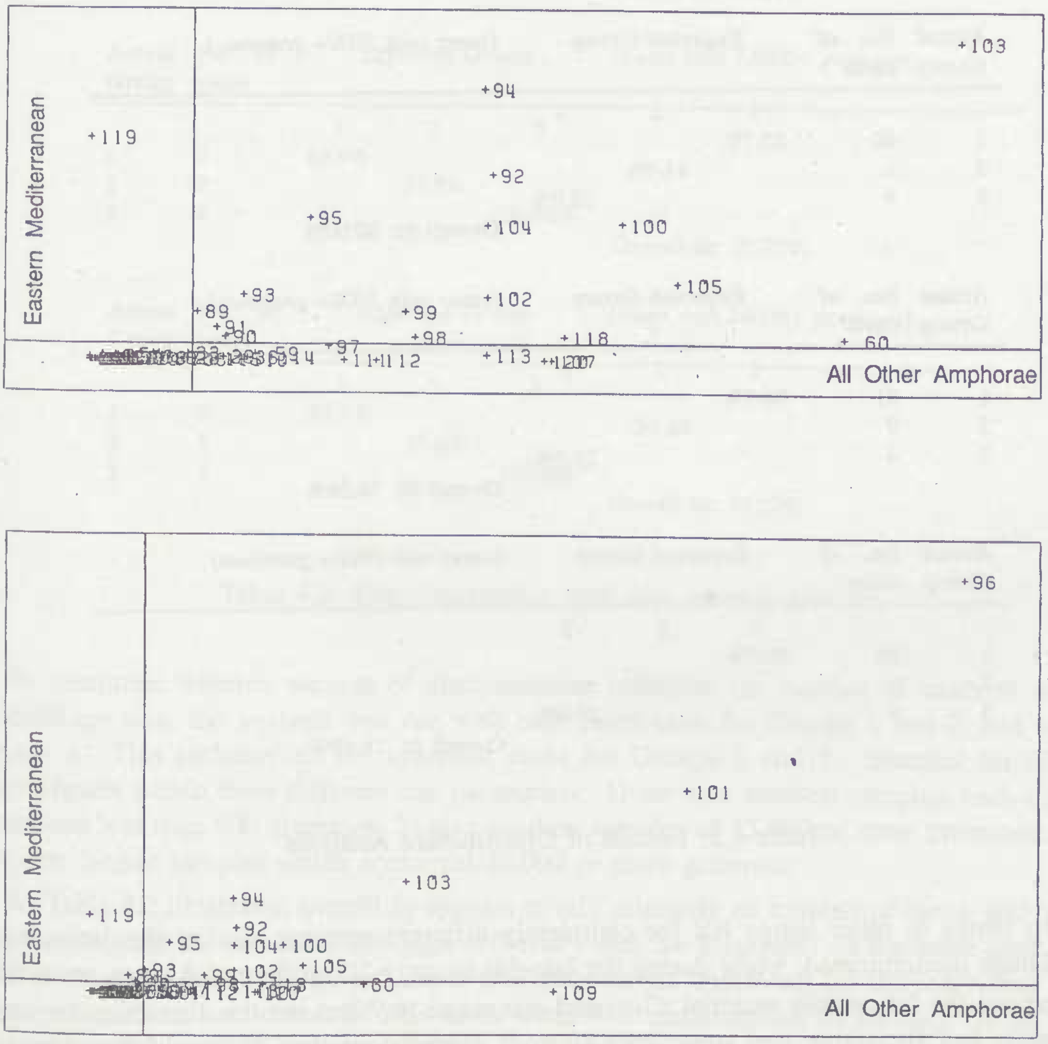

Fig. 4.3: Ratios of eastern Mediterranean amphorae to all other amphorae

Actual Group	No. of cases	Expected Group			(all cases)
		1	2	3	
1	92	91.3%			
2	9		44.4%		
3	4			25.0%	
					Overall fit: 84.76%

Actual Group	No. of cases	Expected Group			(cases with 2000+ grammes)
		1	2	3	
1	62	88.7%			
2	9		44.4%		
3	4			25.0%	
					Overall fit: 80.00%

Actual Group	No. of cases	Expected Group			(cases with 3000+ grammes)
		1	2	3	
1	51	86.3%			
2	9		44.4%		
3	4			25.0%	
					Overall fit: 76.56%

Actual Group	No. of cases	Expected Group			(cases with 5000+ grammes)
		1	2	3	
1	16	89.5%			
2	9		55.6%		
3	4			25.0%	
					Overall fit: 78.43%

Table 4.1: Results of Discriminant Analysis

Red Slip lamps to other lamps but for completely different reasons: during the 1st century Italian lamps predominated, while during the late-4th to early-5th unslipped African ones did.

Taking out the 1st century material alleviated one major problem but the distinction between late Roman and Byzantine was sometimes blurred. Experimentation illustrated that this was related to the number of cases in each group and, to a lesser extent, to absolute assemblage size. Initially Group 1 (late Roman) consisted of 92 cases; Group 2 (Byzantine) of 9; and Group 3 (2nd and 3rd century AD) of 4. Cases in Group 1 were of diverse assemblage size, ranging from less than 100 to 150,000 grammes in total. Analysis was repeated excluding cases from Group 1 on the basis of size so that with each run assemblage size increased as the number of cases decreased.

As illustrated in Table 4.1, there was a gradual decrease in both the fit of Group 1 and overall fit as cases were excluded. However, when analysis was limited to samples of 5,000+ grammes overall allocation, as well as that for Groups 1 and 2, improved. There was not a steady increase after this size level was reached but when all cases in Group 1 were 10,000+ grammes there was at least a 50% success rate for each group.

Actual Group	No. of cases	Expected Group 1	2	3	*(cases with < 500 grammes)*
1	9	100.0%			
2	9		77.8%		
3	4			25.0%	
					Overall fit: 77.27%

Actual Group	No. of cases	Expected Group 1	2	3	*(cases with 13000+ grammes)*
1	9	88.9%			
2	9		77.8%		
3	4			50.0%%	
					Overall fit: 77.27%

Actual Group	No. of cases	Expected Group 1	2	3	*(cases with 20000+ grammes)*
1	9	88.9%			
2	9		55.6%		
3	4			100.0%	
					Overall fit: 77.27%

Table 4.2: Experimentation with case number and size

To determine whether success of discrimination relied on the number of cases or absolute assemblage size, the analysis was run with nine cases each for Groups 1 and 2, and four for Group 3. This included all the available cases for Groups 2 and 3. Samples for Group 1 were chosen within three different size parameters: 1) the nine smallest samples, each of which contained less than 500 grammes; 2) nine random samples of 13,000 or more grammes; and 3) the nine largest samples which contained 20,000 or more grammes.

As Table 4.2 illustrates, overall fit appears to rely primarily on number of cases, and thus the same percentage is calculated although different cases are involved. At the same time, group allocation seems to be related to sample size with results improving for the individual groups when larger cases are included. This same trend was demonstrated by running Group 1 with eighteen samples of different sizes (Table 4.3). Again overall fit is related to the number of cases and when cases of larger sample size are included in the run the allocation of the individual groups improves.

Despite the indication that assemblage size is important, it has not been possible to suggest a minimum weight. This is in keeping with Orton's (Orton 1982, p. 18) premise that it is not absolute quantity that matters, but quantity in relation to the proportion of the particular type.

It was then essential to see if the use of Discriminant Analysis could be successfully used to distinguish between different sites. The methodology was altered since comparable data was not available either in terms of what had been recorded or how this had been done. The field of reference was therefore restricted to published Vandal-Byzantine deposits and the percentages of five eastern Mediterranean amphorae types to the total amphorae were compiled for Benghazi, Carthage and Caesarea. It was expected that these sites would be readily distinguished from

Actual Group	No. of cases	Expected Group 1	2	3 (small cases)
1	18	83.3%		
2	9		77.8%	
3	4			25.0%

Overall fit: 74.19%

Actual Group	No. of cases	Expected Group 1	2	3 (small and large cases)
1	18	77.8%		
2	9		55.6%	
3	4			100.0%

Overall fit: 74.19%

Table 4.3: Results of Discriminant Analysis for eighteen samples of different sizes

each other as they are different distances from the amphorae sources and on different trade routes. The data included weights of rims, bases and handles but not body sherds, as this was not available for all three sites.

Analysis of from two to four assemblages each from Benghazi (Riley 1979), Carthage (Riley 1981, Tomber in press) and Caesarea (Riley 1975) achieved excellent results with 100% discrimination for each group. Furthermore, canonical discriminant functions illustrate that the group centroids are well dispersed with no ambiguity in case allocation.

This was the extent of comparative material using weight data. The same analysis using Benghazi, Carthage and Caesarea was rerun using percentages based on rim, base and handle counts, with equally positive results. The next stage, was to run Discriminant Analysis using different types of quantification, e.g. counts of rims, bases, handles and body sherds versus counts of rims, bases and handles. It was felt that because the overall patterns of long-distance trade reflected by ceramics are very gross, comparison of dissimilar measures would prove successful. To this end mixed count data from additional sites in Italy (Arthur 1985, Whitehouse *et al.* 1985), Sardinia (Villedieu 1984) and Spain (Keay 1984) were added to the core of Benghazi, Carthage and Caesarea. Preliminary results from this analysis are promising, with a high degree of overall discrimination. In addition, the group centroids reflect geographic divisions as one might expect and indicate that Discriminant Analysis will provide a valuable method for inter-site comparison.

4.6 Conclusions

The use of ratios differs very little from traditional applications of percentages for data comparison but they do allow greater flexibility when choosing which aspects of the assemblage are significant for addressing specific questions. Of prime importance is the use of multiple percentages and multiple sites with Discriminant Analysis, for in this way an accurate comparison can be made.

It is suggested that a group of large 'type' deposits be established for the Roman Mediterranean, taking both chronological and geographical diversity into account. New deposits could, in turn, be classified and compared by adding them to the data base and performing Discriminant Analysis. This would also provide detailed information about which variables are most significant for comparison.

In summary, two sources of deposit variability, namely assemblage size and the ratio between types, have been isolated. On this basis it has been suggested how Discriminant Analysis can be useful in comparing deposits when only weight or count data are available.

Acknowledgements

I would like to thank Dr. Stephen Shennan for many long discussions and helpful suggestions; and for reading and commenting on a draft of this paper. I am also grateful to Sebastian Rahtz who, with extreme patience, guided me through the ICL 2976 and enabled me to complete this analysis. Dr. Clive Orton read a draft of this paper after it was presented at the conference and I grateful to him for his comments and suggestions for future work.

References

ARTHUR, P. 1985. 'Naples: notes on the economy of a dark age city', *in* C. Malone & S. Stoddart, (eds.), *Papers in Italian Archaeology IV. The Cambridge Conference Part IV. Papers in Italian and Medieval Archaeology*, pp. 247–59, International Series 246, British Archaeological Reports, Oxford.

DORAN, J. E. & F. R. HODSON 1975. *Mathematics and Computers in Archaeology*, Edinburgh University Press, Edinburgh.

FULFORD, M. G. & D. P. S. PEACOCK 1984. *Excavations at Carthage: The British Mission Vol, I, 2. The Avenue du President Habib Bourguiba, Salammbo. The Pottery and Other Ceramic Objects from the Site*, Department of Prehistory and Archaeology, University of Sheffield, Sheffield.

HAYES, J. 1976. 'Pottery: stratified groups', *in* J. H. Humphrey, (ed.), *Excavations at Carthage 1975 Conducted by the University of Michigan, Vol. 1*, pp. 47–124, Cérès Productions, Tunis.

HAYES, J. & J. RILEY 1978. 'Pottery report—1976', *in* J. H. Humphrey, (ed.), *Excavations at Carthage 1976 Conducted by the University of Michigan, Vol. 4*, pp. 23–98, Kelsey Museum, University of Michigan, Ann Arbor.

KEAY, S. J. 1984. *Late Roman Amphorae in the Western Mediterranean A typology and economic study: the Catalan evidence*, International Series 196, British Archaeological Reports, Oxford.

KEELEY, J. 1986. 'Roman pottery from Cirencester', *in* A. McWhirr, (ed.), *Houses in Roman Cirencester*, pp. 153–189, Alan Sutton Publishing Ltd, Gloucester.

MILLETT, M. 1979. 'How much pottery?', *in* M. Millett, (ed.), *Pottery and the Archaeologist*, pp. 77–80, Occasional Publication 4, Institute of Archaeology, London.

ORTON, C. R. 1975. 'Quantitative pottery studies: some progress, problems and prospects', *Science and Archaeology*, 16, pp. 30–5.

ORTON, C. R. 1982. 'Computer simulation experiments to assess the performance of measures of quantity of pottery', *World Archaeology*, 14 (1), pp. 2–20.

RILEY, J. 1975. 'The pottery from the first session of the excavation in the Caesarea hippodrome', *Bulletin of the American School of Oriental Research,* 218, pp. 25–63.

RILEY, J. 1979. 'The coarse pottery from Benghazi', *in* J. A. Lloyd, (ed.), *Sidi Krebish Excavations, Benghazi (Berenice), Vol. II,* pp. 91–467, Socialist People's Libyan Arab Jamahiriya Secretariat of Education, Department of Antiquities, Tripoli, Libya.

RILEY, J. 1981. 'The pottery from the cisterns 1977.1, 1977.2 and 1977.3', *in* J. H. Humphrey, (ed.), *Excavations at Carthage 1977 Conducted by the University of Michigan, Vol. 6,* pp. 85–124, Kelsey Museum, University of Michigan, Ann Arbor.

SHENNAN, S. J. 1987. *Quantifying Archaeology: an introductory text,* Edinburgh University Press, Edinburgh.

SPSS, 1983. *SPSSX User's Guide,* (manual), London.

THOMAS, C. 1959. 'Imported pottery in dark-age western Britain', *Medieval Archaeology,* III, pp. 98–111.

THOMAS, C. 1981. *A Provisional List of Imported Pottery in Post-Roman Western Britain and Ireland,* Special Report 7, Institute of Cornish Studies.

TOMBER, R. S. 1986. 'Pottery from the south side of the circular harbour', *CEDEC,* 7, pp. 34–58.

TOMBER, R. S. in press. 'Pottery from the 1982 and 1983 Circus excavations', *in* J. H. Humphrey, (ed.), *The Circus and a Byzantine Cemetery at Carthage, I,* Ann Arbor.

VILLEDIEU 1984. *Turris Libisonis. Fouille d'un site romain tardif à Porto Torres, Sardaigne,* International Series 224, British Archaeological Reports, Oxford.

WHITEHOUSE, D. *et al.* 1985. 'The Schola Praeconum ii', *Paper's of the British School at Rome,* LVIII (new series XL), pp. 163–210.

WISHART, D., 1978. *CLUSTAN User Manual Third Edition,* (manual), Edinburgh.