‘Four pots good, two pots bad’: exploring the limits of quantification in the study of archaeological ceramics

Clive Orton, UCL Institute of Archaeology

1. Introduction

This quotation is not to be taken literally, of course. It simply makes the point that there are thresholds, on one side of which quantitative approaches are useful, and on the other side of which they may not be. To explore where these thresholds might lie, I shall take a ‘top-down’ approach, starting from the highest-level or most synthetic questions, and breaking them down into their component parts. I suggest that these questions take the form of inter-assemblage comparisons of their compositions, which could be in terms of ceramic fabrics (as a proxy for source), form (as a proxy for function), ‘type’ (as a proxy for date) or of other variables such as size or capacity (this list is not intended to be exhaustive). This is not to decry the value of the study of individual vessels or types, but to argue that these are building blocks in a bigger structure – the study of assemblages in their context. This, I believe, is what distinguishes archaeology from art history, antiquarianism or collecting.

The reasons for choosing this level of study of assemblages are taphonomic. Lower-level studies of assemblages do not seem to be useful, e.g.

(1) the overall size of an assemblage (however it is measured): to mean anything, this must relate to a ‘complete’ assemblage (otherwise we don’t know how much is missing), implying total excavation and recording, and how often do we see that?
(2) the composition of a single assemblage: how can we interpret a single composition? Without a basis for comparison it does not really tell us anything.

A further point is that the assemblages we handle and study are usually only a fraction, or a sample, of the pottery that was actually in use at a certain point at a certain time, and that our prime interest should be in these original ‘life’ assemblages rather than in the ‘death’ assemblages that we actually have to hand. Of course, the latter are all that we have, but they should be seen as a pointer towards the former rather than as objects of study in their own right. There will also be cases when the taphonomic processes themselves are of interest (e.g. Orton 1985) (Section 2.6). This leads us into difficult areas of sampling theory and estimation, which for many archaeologists are dangerous *terrae incognitae*, but which must be explored if we are to make progress.

But first, to explore the relative roles of quantitative and qualitative approaches, we need to investigate the sorts of data that go to make up our assemblage compositions; where they come from and how we use them.

2. Sources of data

As suggested above, a basic set of ceramic data might consist of:
Fabric, form, date, size, to which must be added decoration, and possibly other variables. In addition, some measure of quantity must be included if we are to estimate the

compositions of assemblages. These summative variables are themselves built up of more specific variables, some of which may be more objective and others more subjective, but none the less valuable. Individual knowledge and experience is valuable here – a key issue in British archaeology is how this expertise can be transmitted from one generation to the next. This question is particularly apposite at a time when many of the specialists are dying or reaching retirement age. It's worth looking at each area in more depth.

2.1 Fabric data

The assigning of sherds to fabrics, or to wares, has long been regarded as a subject for the exercise of individual expertise or scholarship, and rightly so. The ability to recognise fabrics is something that comes from a lifetime of experience, together with mentoring from established specialists and discussions with colleagues. Attempts to replicate such skills in explicitly quantitative ways, e.g. by the use of expert systems, have not been successful (Wilcock 1999, 38). And yet, we still have a problem. Expertise is a rare commodity, which needs to be shared, especially in the present environment of commercial archaeology post-1990 and the period of expansion (1975–1990) that preceded it. Quite simply, there were not enough specialists to go round, and bottlenecks built up. As mentioned above, this situation is now exacerbated in the UK because many current specialists are reaching the ends of their working careers.

How is their collective knowledge and experience to be disseminated more widely, and passed on to future generations? Traditionally, this has been done through the medium of written descriptions, but anyone who has tried to match a sherd in their hand with a description on the page will appreciate the weakness of this approach. There often seem to be key characteristics for which one struggles to find words. For example, a certain ware (VRW in the London terminology) can be characterised by the nature of its fracture, which I called "roches moutonnées": I know what I mean, and I can use it, but could anyone else? A distinct improvement is offered through the use of high-quality colour images of surfaces and fractures of sherds of known fabrics, as exemplified by Tomber and Dore (1998), and very useful it is too. There is still the drawback to this approach is better at telling the reader what X ware looks like, than it is at identifying the sherd in their hand. An alternative approach, developed in the 1970s to meet this specific need, is the so-called *Polstore* system at the Museum of London (named after the brand of cabinet in which it is stored, see Orton 1979; Rhodes 1979). Here the user is asked various questions about their problem sherd, the answers to which guide them to a particular part of the storage system, where they will find sherds (possibly of many different wares) which match the physical description of their sherd. From then on, identification is a matter of subjective matching and judgement. In practice, regular users find that the system trains them, and that the more they do, the less they actually need it. Here we see quantitative and qualitative approaches working hand in hand. The problem is that one needs physical access to the storage system (currently located in a not very central part of London) to be able to use it. This leads to the idea of the electronic reference collection (eRC) (Lange 2004) in which the process is carried out on web-based software, with high-quality images replace the physical sherds (which would still exist somewhere, but not necessarily all in the same place, as the ultimate reference), thus opening up access to

anyone with a web browser. This idea currently languishes for lack of funding – to be really successful it needs to be done on an international basis.

Detailed sourcing of wares can only be conclusively confirmed through scientific (i.e. elemental) analyses, using techniques such as NAA since the 1950s (Sayre and Dobson 1957) and ICP-AES since the 1980s (Hart and Adams 1983). Often, several sources in the same region may have produced wares that to the naked eye and even microscopically are very similar, and if identification to the exact source (rather than to the generic region) is needed, then more sophisticated approaches must be used. Such techniques are inevitably expensive, so that it will only ever be possible to apply them to a very small proportion of excavated material. Well-thought-out programmes of sampling are needed to ensure that best use is made of such costly resources.

2.2 Form data

Every sherd has a fabric, but not every sherd has a recognisable form (Darling 1989). Nevertheless, assignation of sherds to forms is another crucial and widespread archaeological activity. Here expertise may show itself in the ability to ascertain the form of ever-smaller sherds, and once again this has traditionally been regarded as a subjective qualitative skill. Typically, forms are recognised at two levels – a broad general level (e.g. jar, jug, bowl … (MPRG 1998), which interestingly does not translate well from one language to another) which may be taken to indicate function, and a detailed typological level which may relate more to source and/or chronology. The latter often breaks down into a plethora of sub-levels, which may tell use more about the psyche of the researcher than the nature of the pottery.

Just as for fabrics, early quantitative attempts at characterising and classifying forms were not particularly successful (e.g. Wilcock and Shennan 1975). Cluster analysis in particular seems to have been rather a blunt instrument (Baxter 2003, 92; 2008, 972). However, there is evidence that things are now changing, with increasingly sophisticated ways of capturing, comparing and classifying ceramic profiles. Simple measurements and ratios proved useful in classifying rather amorphous hand-made vessels (Richards 1987, 71–76), and the introduction of the tangent-profile method suggested a way ahead (Leese and Main 1983). More recently, advanced scanning technology and profile capture is beginning to make an impact (Karasik and Smilansky 2008) although the equipment is expensive and needs a vast throughput to justify the capital cost. Nevertheless, it has been possible to distinguish successfully between very similar products from different sources and to correlate such shape differences with differences in fabrics (Adan-Bayewitz *et al*, forthcoming). The technology may actually be better than the human eye in such repetitive tasks as measuring diameters (Section 2.4) and ascertaining the angle of orientation of sherds, thus aiding drawing and reconstruction.

Related to questions of form are those related to decoration, because both express aspects of shape, but on different scales. Particularly for highly-decorated wares, such as samian, this has long been the preserve of the experienced specialist, on whom we have all come to rely. To take the example of decorated samian, it would be impossible for the average

Roman pottery specialist to develop and maintain the necessary expertise – they would have no time for anything else – so recourse to a small band of dedicated specialists is essential. Reference books (for example Webster 1996) are of course useful, but as we have seen above, the information tends to flow the wrong way, from the general to the particular. This still leaves us vulnerable to 'slings and arrows' – the ageing specialist population and the problems of handing on knowledge to a new generation. Ceramic decoration appeared to be an early candidate for the support of expert systems, which could guide the user from sherd to attribution, but attempts were unsuccessful. It may be that when the capture and comparison of forms has become commonplace, attention will again be turned to the question of decorative motifs, although I suspect that this will be a more difficult nut to crack.

2.3 Chronological data

It is rare for ceramics to carry intrinsic means of dating, although a few instances are known (e.g. C_{14} dating of charred organic inclusions, see Evans and Meggers 1962; Glover 1990). Nevertheless, pottery is widely seen as one of the principle sources of archaeological dating evidence (too widely, some ceramic specialists might add). The evidence is thus usually secondary, being based on typological studies of form and decoration (and to a lesser extent, fabric). This, too, is regarded as an area for the exercise of specialised expertise and accumulated scholarship.

Paradoxically, this was one of the first areas of ceramic studies to which quantitative methods were applied, with Petrie's invention of seriation (Petrie 1899). Although much misused, this has remained a useful tool, provided its basic premises are observed. Obstacles to its uncritical use concern the integrity of assemblages, particularly in relation to residual and infiltrated finds (Carver 1985). Such issues, together with the question of the lifespan of 'ordinary' ceramics (even if the 'heirloom' problem is excluded), also bedevil more traditional qualitative approaches to dating frameworks (Orton and Orton 1975). The way ahead, for both qualitative and quantitative approaches, may lie in more detailed theoretical discussion of the relationship between assemblages and their contexts (e.g. Roskams 1992; Berry 2008), as well as agreement on apparently simple questions like 'what does the date of a ceramic type mean: date of manufacture? Date of contexts in which it is commonly found?' The need for a common definitional framework is apparent from both perspectives.

2.4 size data

Here, by contrast, we are firmly in the realm of quantitative data. The size (diameter) of rim and base sherds can easily be measured by the use of a simple chart (Orton *et al* 1993, 172–3) and, with more difficulty, body sherds can be measured by templates. This is not, of course, an exercise in collecting data 'because it's there', but an input into issues of capacity and hence functions. For example, trends in the diameter of open vessels through the Roman period in north Africa suggest changes in eating practices during that period (Hawthorne 1996), and more subtle changes may reflect changes in basic units of measurement (Thorne, *pers comm*). Measuring (e.g.) rim diameter is also a useful step towards quantifying assemblages (Section 2.5).

2.5 measuring quantities

This quick tour of the different types of data brings us back to my stated aim of characterising and comparing the compositions of assemblages. What is needed is an agreed measure of the quantities of different types of pottery in an assemblage, which can lead to reliable inter-assemblage comparisons, by minimising the risk of bias from factors that do not relate to the original 'life' assemblages (e.g. relative breakage rates, see below). The main contenders are: sherd count, sherd weight, sherd volume/area (*n.b*. not vessel capacity), vessels represented (minimum, maximum or estimated) and estimated vessel equivalents ('eves'). Their relative merits have been much discussed (see Orton *et al* 1993, 166–181 and references there), and this is not the place to repeat the arguments. But I will say that I do not regard any of these, even in combination, as the 'last word'; the future may well lie in the Monte Carlo simulation (Hammersley and Handscomb 1964) of the creation of excavated assemblages, as a means of comparing their parent 'life' assemblages.

2.6 'spin-off' and taphonomic data

We have already touched upon the importance of the relationship between excavated assemblages and their parent 'life' assemblages. Attempts to explore this relationship have led to the discovery of ways of deriving information about the processes involved, which turn out to be of considerable interest in their own right. Derived statistics such as completeness and brokenness (Orton 1985) can augment more traditional approaches, such as the degree of wear (Needham and Sørensen 1989) and the analysis of cross-joins. That the analysis of cross-joins could make a valuable contribution to site analysis is beyond doubt (e.g. Moorhouse 1986 and the Farthest Migrant Matrix Score, see Berry 2008, 92), but it seems to have been held back by a lack of theory. The time and space required to achieve reliable measures of the extent and nature of cross-joins on a site add to the difficulties of using this approach. It also takes a mental adjustment for the ceramic specialist to think of sherds as indicators of the movement and dispersal of archaeological deposits, as well as objects in their own right.

3. synthesis

We have so far seen the greater or lesser extent to which quantitative thinking and practice have permeated the various 'building blocks' from which we must construct our archaeological stories. But "Science is built up of facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house" (Poincaré, 1902). We now have to pull together and integrate all these various strands of evidence (fabrics, forms, decoration, etc.). Here we need both theory and methodology. The theory will come from archaeological expectations and possibilities (e.g. about modes of production, see Peacock 1982, 8–11, or means of distribution, see Fulford and Hodder 1974), while the methodology will comprise the tools needed to link these to the data. As suggested at the start of this paper, these tools will include the comparison of assemblage compositions, perhaps chronologically or perhaps spatially (or indeed both), to answer questions about (for example) chronological trends of geographical distributions (e.g. Going 1987, figs 52–59). When inter-site comparisons are involved, it is likely that the

data will have been collated from a variety of sources, compiled by different people, perhaps over a wide time-span. This brings additional difficulties, as definitions and measures may well differ from one source to another, perhaps reflecting growing knowledge of particular wares, or perhaps to different levels of resources available at different places and different times. This raises the thorny issue of standardisation of terminology, without which quantitative studies can be impossible, or reduced to a 'lowest common denominator', but which is difficult to achieve and within which it is difficult to accommodate advances in knowledge.

One question that often arises is "is there a minimum size of assemblage below which it is not worth quantifying?" i.e. is there a *threshold*? There is a valid point behind this question – some samples (whether of ceramics or anything else) are simply too small to yield reliable information, and if they are used they can be misleading. But where should one draw the line? It depends not only on the size of the assemblage, but on the questions that are being asked of it, and we have to admit that our current questions are not the only ones that could be asked. Further, to know on which side of the line an assemblage sits, we need to carry out at least a basic quantification, which it would then seem perverse to ignore. Finally, it is very often possible to aggregate assemblages in different ways for different questions, so that even a very small assemblage may be able to contribute to a useful statistic. For all these reasons, I resist the idea of imposing a threshold on the quantification of ceramic assemblages.

More important than questions of size are issues of the integrity and status of assemblages and their contexts. The major problem of residuality and the lesser one of infiltrated finds have long been recognised, but surprisingly little has been done on them. After a wide-ranging study, Berry (2008, 232) suggests that the effects of residual and infiltrated finds may be less than is sometimes feared, but this conclusion must be seen in the context of his aim of characterising ceramic assemblages in order to assist in the interpretation of the deposits in which they were found. Their effect on more mainstream ceramic interpretations (e.g. dating or distribution) remains a worry, on which relatively little work has been done, considering its perceived importance.

4. conclusion

A quantitative framework for analysing and comparing data is essential in ceramic studies, even if the data themselves are from a variety of sources, some of which may be qualitative or even quite subjective. Beyond this, quantitative approaches have a valuable function in retaining and transmitting knowledge, and increasingly in the automation of basic repetitive tasks, thus freeing up the archaeologist's time for more important and interesting tasks such as thinking about what the data might mean. Modern technology can replace the myriad of small 'local' decisions (such as the diameter and angle of a rim sherd) which can clog up and weary our minds. The sophistication of modern software is leading us towards the convergence of quantitative and qualitative data (e.g. in the shape and decoration of vessels), to the point where we may not be able, or need, to distinguish between the two. Some may see this mechanisation as 'dumbing down' or as usurpation

of the role of the specialist; I see it more as freeing the specialist from the routine to spend more time on the really interesting aspects of the subject.

Bibliography

Adan-Bayewitz, D., Karasik, A., Smilansky, U., Asaro, F., Giauque, R.D. and Lavidor, R. forthcoming. Differentiation of ceramic chemical element composition and vessel morphology at a pottery production center in Roman Galilee. Submitted to *Journal of Archaeological Science*.

Baxter, M.G. 2003. *Statistics in Archaeology*. London: Arnold.

Baxter, M.G. 2008. 'Mathematics, statistics and archaeometry: the past 50 years or so' *Archaeometry* 50, 968–982.

Berry, M.G. 2008. *Stratigraphic and material interpretation of site evidence: Investigations toward the nature of archaeological deposits*. Unpublished Ph.D. thesis, University of York.

Carver, M.O.H. 1985. 'Theory and practice in urban pottery seriation' *Journal of Archaeological Science* 12 (5) 353–366.

Darling, M.J. 1989. 'Nice fabric, pity about the form' *Journal of Roman Pottery Studies* 2, 98–101.

Evans, C. and Meggers, B.J. 1962. 'The use of organic temper for Carbon 14 dating in lowland South America' *American Antiquity* 28, 243–245.

Fulford, M.G. and Hodder, I. 1974. 'A regression analysis of some later Romano-British pottery: a case study' *Oxoniensia* 39, 26–33.

Glover, I.C. 1990. 'Ban Don Ta Phet 1984–85' in Glover, I.C. and Glover, E. (eds) *Southeast Asian Archaeology 1986*. British Archaeological Reports, International Series 561 (Oxford: BAR) 139–183.

Going, C.J. 1987. *The Mansio and other sites in south-eastern sector of Caesaromagus: the Roman pottery*. CBA Res. Rep. 62

Hammersley, J.M. and Handscomb, D.C. 1964. *Monte Carlo Methods*. London: Chapman and Hall.

Hart, F.A. and Adams, S.J. 1983. 'The chemical analysis of Romano-British pottery from the Alice Holt Forest, Hampshire, by means of inductively coupled plasma emission spectrometry' *Archaeometry* 25, 197–285.

Hawthorne, J. 1996. Commensalism and common sense: a new approach to archaeological ceramics. *Assemblage* 1 http://www.assemblage.group.shef.ac.uk/1/hawth.html, visited 19/01/2009.

Karasik, A., and Smilansky, U., 2008. 3D scanning technology as a standard archaeological tool for pottery analysis: practice and theory. *Journal of Archaeological Science* 35, 1148–1168.

Lange, A.G. (ed.) 2004. *Reference Collections. Foundations for Future Archaeology*. ROB, Amersfoort.

Leese, M.N. and Main, P.L. 1983. 'An approach to the assessment of article dimensions as descriptors of shape' in Haigh, J.G.P. (ed.) *Computer applications and Quantitative Methods in Archaeology 1983*, 171–180.

Moorhouse, S. 1986. 'Non-dating uses of medieval pottery' *Medieval Ceramics* 10, 85–124.

MPRG 1998. *A Guide to the Classification of Medieval Ceramic forms*. MPRG Occasional Paper 1.

Needham, S.P. and Sørensen, M.L.S. 1989. 'Runnymede refuse tip – a consideration of midden deposits and their formation' in Barrett, J.C. and Kinnes, I.A. (eds) *The archaeology of context: recent research on the Neolithic and Bronze Age in Britain*, 113–120.

Orton, C. 1979. 'Dealing with the pottery from a 600-acre urban site' in Millett, M. (ed.) *Pottery and the Archaeologist*, Institute of Archaeology Occasional Papers 4, 61–71.

Orton, C. 1985. 'Two useful parameters for pottery research' in Webb, E. (ed.) *Computer Applications in Archaeology 1985*, 114–120.

Orton, C. and Orton, J. 1975. 'It's later than you think: a statistical look at an archaeological problem' *London Archaeologist* 2 (11), 285–287.

Orton, C., Tyers, P. and Vince, A. 1993. *Pottery in archaeology*. Cambridge Manuals in Archaeology.

Peacock, D.P.S. 1982. *Pottery in the Roman World: an ethnological approach*. London: Longman.

Petrie, W.M.F. 1899. 'Sequences in prehistoric remains' *Journal of the Royal Anthropological Institute* 29, 295–301.

Poincaré, J.H. 1902. *La Science et l'hypothèse*.

Rhodes, M. 1979. 'Methods of cataloguing pottery in Inner London: an historical outline' *Medieval Ceramics* 3, 81–108.

Richards, J. 1987. *The significance of form and decoration of Anglo-Saxon cremation urns*, British Archaeological Reports, British Series 166.

Roskams, S. 1992. 'Finds context and deposit status' in Steane, K. (ed.) *Interpreting Stratigraphy* 1, 27–9.

Sayre, E.V. and Dobson, R.W. 1957. 'Neutron activity study of Mediterranean potsherds' *American Journal of Archaeology* 61, 35–41.

Tomber, R. and Dore, J. 1998. *The National Roman Fabric Reference Collection : a handbook*. Museum of London Archaeology Service.

Webster, P.V. 1996. *Roman samian ware in Britain*. Council for British Archaeology, Practical Handbook, York.

Wilcock, J.D. 1999. 'Twenty five years of statistical and other techniques' in Dingwall, L., Exon, S., Gaffney, V., Laflin, S. and van Leusen, M (eds) *Archaeology in the Age of the Internet*. BAR International Series 750, 35–51.

Wilcock, J.D. and Shennan. S.J. 1975. Computer analysis of pottery shapes. *Computer Applications in Archaeology 1975*, 00–00.

http://www.weizmann.ac.il/complex/uzy/archaeomath/# (visited 02/02/09)