in: S. Verdan, T. Theurillat and A. Kenzelmann Pfyffer, eds., Early Iron Age Pottery: A Quantitative Approach (BAR-IS 2254), Oxford 2011, pp. 23-38.

'ERFAHRUNGSBERICHT' OF APPLICATION OF DIFFERENT QUANTITATIVE METHODS AT KALAPODI

Sara STRACK

Abstract

Late Bronze and Early Iron Age levels in a sample trench at the Central Greek cult site at Kalapodi are used to apply a range of quantitative methods, in order to assess their comparability, as well as applicability and ease of use for an assemblage characterized by a high degree of fragmentation and corresponding low completeness. Sherd count, weight, rim and base portion EVEs, MNIs, and aggregate feature count are assessed for their uses and pitfalls. While none of these methods are found to be entirely unproblematic, and while all have their advocates and opponents in the relevant bibliography, it is argued that particularly for the analysis of pottery as evidence for human behaviour, but also for the comparison of assemblages across sites and/or periods, methods establishing vessel numbers (EVE, MNI, AFC) are preferable over those establishing the amount of pottery (sherd count, weight). A comparison of data sets resulting from different quantitative methods illustrates that the latter are roughly equivalent, suggesting that the quantitative approach best suited for each assemblage can be chosen while retaining inter-site and inter-period comparability. Finally, an examination of sample size effect on representation of artefact classes shows the need for sample sizes larger than anticipated in order to obtain statistically significant data.

Keywords: Kalapodi, Phokis, EVE, MNI, aggregate feature count, methodology, sample size effect

INTRODUCTION

The use of a quantitative approach for the study of material remains has a long history in archaeology (see Orton 1993, 169–177, for an overview of history and methodology). While initially, quantification largely focussed on seriation and typo-chronology (see references in Orton 1993), in recent years it has been used to study exchange and the volume of trade (e.g. Tomber 1993); the magnitude and organization of ceramic production (e.g. Shaw *et al.* 2001, especially van de Moortel's contribution; Knappett 2001); as well as, more generally, ceramic use and discard behaviour. While thus useful for a range of research questions and a variety of assemblages, quantitative methods are of particular value for the study of large sherd assemblages, which usually consist of fragmentary, and often frankly unappealing, material deriving from dumps and fills. While not necessarily adding to the art historical account of their period, such deposits can provide a glimpse into the day-to-day activities of a site's occupants and thus represent a valuable resource for the understanding of past human behaviour.

In preparation for the workshop 'Early Iron Age pottery: a quantitative approach', the author and her collaborators focussed on studying pottery from one sample trench (Kalapodi 5030/4965). Keeping in mind the overarching theme of the workshop, we applied a range of quantitative methods in order to assess their applicability and usefulness for our material and gathered empirical data relating to time and effort required for each method in relation to its results. In the following, the results from this pilot study are discussed, together with an attempt to assess the validity of the different methods applied.

The find situation at Kalapodi is outlined in greater detail elsewhere in this volume (Kaiser, Rizzotto and Strack); the main characteristics, however, are here summarized again. The vast majority of finds from the site derive from fill strata, probably brought in during terracing operations (e.g. Felsch 1987, 5). In the sample trench processed for this workshop, neither closed nor primary deposits (i.e., in situ depositions of material) could be identified. The range of dates present in each of the seven stratigraphic units (SU) and four chronological phases supports the notion that the material contained in these SUs accumulated elsewhere on site and was moved when terracing of the hill slope required large amounts of soil or fill.

The sherd size and nature of breaks indicated to us that pottery at Kalapodi was very likely broken intentionally after use. This fact contributes largely to the typical aspect of sherd material at the size – on the one hand, sherds are well preserved and usually easily identifiable; on the other hand, brokenness is high with a low corresponding completeness, resulting in rare occurrences of complete or largely complete vessel profiles.

Ceramics used at Late Bronze and Early Iron Age Kalapodi were, for the most part, decorated very simply, with monochrome painted bodies and monochrome or banded vessel rims predominating. A second major characteristic is the high frequency of undecorated, handmade coarsewares, including cooking wares, which represent the majority of medium and large closed vessels in the Early Iron Age. Previous publications of material from Kalapodi clearly reflect the fragmentary nature of finds, with catalogues dominated by single sherds (Jacob-Felsch 1987, Nitsche 1987, Jacob-Felsch 1996). The nature of ceramic finds at Kalapodi should thus resemble the situation at other sanctuaries and, in particular, settlement sites, where much of the excavated material derives from continuous deposits and essentially represents rubbish discarded by the occupants. The mixture of decorated and plain, table and utilitarian wares at Kalapodi, too, is mirrored in the ceramic assemblage especially by deposits of domestic nature (e.g. Tsoungiza: Thomas 2005; Asine: Wells 1983). Conversely, finds from burials stand in stark contrast, since here complete vessels are usually preserved and selective deposition favoured pieces with lavish decoration or of unusual shape or function.

Given the fragmentary nature of the evidence, a precise typological identification of sherds was often not possible beyond the broad classes outlined in the introduction to the sorting system at Kalapodi (Fig. 1; Kaiser, Rizzotto and Strack, this volume). Thus for example, rim fragments from cups and skyphoi often have similar profiles and decoration, and while the average rim diameter of skyphoi appears to be larger than that of cups, there is enough overlap between the two shapes to render an accurate distinction based on small rim fragments impractical. Similar overlap in shape and size exists between different types of medium and large closed shapes (e.g. amphora, hydria, jug) and handmade coarsewares (e.g. amphora, jug, jar; one- and two-handled cooking pot). We expect that further study of the pottery from Kalapodi will allow us to establish a site-specific typology that will aid in better distinguishing between different types among the fragmentary material. Consequently, our study of pottery from Kalapodi initially focused on functional classes, rather than individual shapes and types. Thus, we differentiate between vessels for serving and consumption (small open shapes – cups, skyphoi, kylikes, etc.; also large open shapes – kraters), transport and storage (medium and large closed shapes – amphorae, hydriae, jugs), bulk storage (large and very large closed, usually coarse – jars, pithoi), and cooking (cooking pots), in order to illuminate human behaviour at the site.[1]

It is with the above in mind that the findings from trench Kalapodi 5030/4965 are here presented as a case study for the application of quantification and the mathematical manoeuvres possible on the basis of sherd counts.

DIFFERING RESEARCH STRATEGIES BEHIND VARIOUS QUANTITATIVE METHODS

The methods used for analyzing our sample can be distinguished into those resulting in simple quantification (how much pottery) and those assessing numbers of vessels (how many pots) (e.g. Orton 1982, 1). Among the former are sherd count and weight; the latter include estimated vessel equivalents (EVE) based on rims and/or bases, and minimum number of individuals (MNI). A variant of the sherd count, based on feature sherds (rims, handles, bases), is here referred to as aggregate feature count (AFC) and was found to yield results closer to the EVE and MNI values than the simple sherd counts (for more discussion of the AFC, see *infra*).

In assessing the usefulness of these various means of pottery quantification, and in evaluating their applicability for archaeological excavations which often impose constraints in time, space, and personnel, it should be noted that the two basic strategies of quantification (bulk and number) pursue different results. Sherd count and weight allow for a basic record of the amount of pottery present; all sherds from a given context can be accounted for and be recorded in as high a level of details as desired; and last but not least, because of their low demands on time and equipment, count and weight are broadly applied and thus allow for inter-site comparison of datasets. When converted into relative figures, a cautious assessment can be made of representation of vessel classes; however, comparative representation based on sherd count and/or weight does not reflect the actual representation of vessels of different ceramic classes and can be misleading (see, for example, the figures for handmade coarseware at Kalapodi, *infra*). Thus, sherd count and weight are biased toward vessels with large surface area and/or high breakage rate (count), or large size, heavy fabric, and/or thick walls (weight) (see Fig. 5 for estimates of sherds per vessel unit, Fig. 4 for average weight of individual sherds, and Fig. 3 for the share of features sherds of the overall count, for different ceramic categories).[2]

Methods recording vessel numbers, conversely, allow for an investigation of the composition of a given assemblage and thus are instrumental in understanding ceramic use as reflected in these remains (for the effect of vessel life-span

[1] The sorting concept is built on these distinctions and works on the assumption that a fair number of sherds are no further identifiable than these categories. For the most part, the same should be true for sites even with better preservation and better known typology, since featureless body sherds usually cannot be classified any further than this. The idea behind the sorting concept is to allow inclusion of all sherds found in the initial analysis. Identification of types and shapes subsequently takes place on the basis of a selection of sherds with more closely identifiable characteristics. Note that there are functional classes beyond the ones listed above – for example, there are vessels for food preparation other than cooking (e.g. coarseware basins and other, medium and large open shapes clearly not used for storage), pouring of liquids (trefoil jugs, both fine and coarse, which are presumably not primarily used for storage and/or transport of liquids), containers for special liquids (small jugs – function here not entirely clear, since they could be used for pouring liquids in small amounts, but also for perfumes and unguents – aryballoi – or the dedication of container and liquid). These classes are rather sparsely represented and are largely left aside for the purpose of this initial discussion. For future, more detailed studies of the material, these classes obviously warrant closer examination.
It should be pointed out that the sorting system employed here aims at the identification of vessel function and use, and that applications such as seriation based on typology of fabric classes is neither an aim, nor a likely outcome of our initial approach; the basic categories identified are ubiquitous in all levels discussed here and continue beyond the chronological limits of our study at Kalapodi. At the same time, the recording of more details regarding fabric and shape, as outlined by Kaiser, Rizzotto and Strack elsewhere in this volume, will make such studies possible in the future.

[2] Wall thickness, hardness, and fabric cohesion, as well as manufacturing technique, influence breakage patterns; consequently, sherd count skews towards different vessel classes in different assemblages. In Late Bronze and Early Iron Age assemblages studied by the author, medium and large shapes tend to break into large sherds due to their relatively thick walls, whereas small open shapes break into a multitude of small, thin fragments. Handmade coarsewares, on the one hand, are often thick-walled and thus break into large pieces, while on the other hand, low firing temperature, the presence of large amounts of coarse inclusions, and at times the lack of homogeneity of the fabric often lead to crumbling. Depending on the preservation of pottery at a given site, these factors can determine the possibility of identification of certain classes to a notable degree. The presence of pithos sherds tends to dominate sherd weight, and thus can impede comparison of weight in deposits with differing amounts of pithos ware.

on assemblage composition, see e.g. Orton *et al.* 1993, 166–167; Schiffer 1987, 5–7; for thoughts on the situation at a sanctuary site, see Kaiser, Rizzotto and Strack, this volume). When applied to comparative studies of assemblages, numbers converted into relative figures (percentages) should not differ greatly between the different methods available (see *infra* for discussion of comparability); however, the applicability of different methods to the assemblage at hand varies, both in the literature and our own experience.

It should be noted that outside Mediterranean archaeology, statistical methods of assemblage analysis are used besides simple quantification as presented here. Thus, there are numerous studies which use diversity, i.e., the number of artefact categories present and their relative distribution, as a tool for assemblage characterization (Jones and Leonard 1989; Kintigh 1989, 25–26). A comparative assessment of methods based on quantification, proportional frequency and diversity by Cannon (1985) concludes that diversity is best suited for investigating function and behavioural patterns since it 'is not so encumbered by the effects of extraneous factors or unjustified assumptions as are the other measures' (Cannon 1985, 791; also Kintigh 1984; for a broader assessment of diversity and its uses for quantification, see e.g. Bobrowsky and Ball 1989). At present, the use of descriptive statistics is prevalent in the Aegean and beyond; the establishment of better site typologies, a deeper understanding of the interrelation between material culture patterning and human behaviour, artefact studies across the boundaries of material (metal, ceramic, stone, etc.), and more standardized artefact typologies might allow the profitable use of more elaborate statistical tools in the future.

Lastly, it is important to note that when adhering to a small number of common-sense principles, and with the requisite amount of transparency regarding sample provenance, size, selection, and methodology, the kind of quantitative approach used does not matter nearly as much for the usefulness and inter-site comparability of the resultant data as does the underlying concept of classification. The final stages of the Bronze and the Early Iron Age see a great degree of regional variability in pottery styles and production, rendering the formulation of a standardized, Aegean-wide typology at any level of detail somewhat dubious; it is for this reason that authors should err on the side of more information in including in publications as much data as possible regarding the definition of ceramic categories and their numerical representation.

DISCUSSION AND ASSESSMENT OF THE RANGE OF METHODS APPLIED AT KALAPODI

Since the main aim of the present paper is to assess validity and usefulness of different quantitative methods, and dependent calculations, all pottery sherds from the sample trench have here been used as the basic dataset; no differentiations have been made by context or phase, mainly in order to maximise sample size.[3] It is hypothesized that material from this trench represents a random sample of ceramics from the Late Helladic IIIC through Late Geometric strata from the site; there is no obvious indication among the finds that would suggest otherwise (e.g., a cache of votive vessels, etc.; but see, on sample size and validity, Leonard 1997, 714; O'Neil 1993). The resulting sample comprises 10,526 sherds, weighing 133.158kg (for a breakdown of this figure by vessel category and sherd type, see Fig. 2). Note that factors such as sample size, brokenness, and representation of rare ceramic classes all contributed to the degree of usefulness of different methods for the material under study.

Sherd count and weight

Counting sherds is possibly the easiest method of pottery quantification available, if only because no special tools are needed to conduct a count. Counting should be done before mending, in particular if average sherd size is to be determined using count and weight; an exception to this rule are fresh breaks (i.e. breakage that occurred during excavation or post-excavation processing). At Kalapodi, sherds are recorded in categories as detailed as possible; lumping together the resulting data is always possible, although information not recorded then cannot be retrieved later. The initial count thus can take the form of a running commentary on the assemblage under study. For example, SUs I and Ia in the Kalapodi sample trench contained a range of cooking pot fabrics, among which we could distinguish at least four distinct categories (wheelmade thick-walled, wheelmade thin-walled, handmade local/regional, and handmade non-local); however, for anything other than a study focussing on cooking pot types, these categories are summarized under handmade and wheelmade cooking wares.

In our experience, weighing is most conveniently done immediately after counting (in fact, sherds are counted into a bowl or bucket which is then weighed for every category). Among the methods assessed here, weighing of sherds was by far the fastest. Generally, the conclusion reached by Gifford (1951, 223) seems correct: 'Weighing sherds seems to give more accurate statistical results than counting them. Accident or the character of the past may account for the size of, and consequently the number of, sherds in a given block excavated, so that the count is variable. But no matter how large or how small the sherds are in a given block, their total weight will remain the same'; similar conclusions are reached by Solheim (1960, 325), Orton *et al.* (1993, 169), and Byrd and Owens (1997, 316).[4]

[3] Unless otherwise indicated, the following figures are based on all lots in trench 5030/4965, with the exception of six lots from which sherds had been discarded before the material could be studied (lots KAL05.13, .21, .26, .167, .173, .177).

[4] Note that Slane 2003, 324, compared sherd weight immediately after initial processing (i.e., washing and drying for 2–3 days) and 4 to 6 weeks after excavation, and found significant differences – apparently dry sherds still contained enough moisture to result in a 8–12% difference in the two weighings. Incomplete drying affects thick-walled and coarse wares to a greater degree than thin-walled and fine wares.

The combination of counts and weights proved very useful to assess the degree of breakage, by establishing the average sherd weight, which can be used to investigate the depositional history of different assemblages (e.g. Rutter 1990, 378–379), as well as typical breakage patterns for different vessel classes (Solheim 1960 examines both aspects).

However, unless the depositional history of different vessel categories within one SU is arguably divergent, this effect should not noticeably change the results within one deposit; the conversion into relative figures eliminates discrepancies between different deposits given that the comparative breakage factor within a deposit is linear (i.e., each ceramic category has an identifiable and stable breakage pattern in relation to other categories). At the same time, when applied to inter-site comparison of assemblages, or study of assemblages across periods, simple counts and weights can be insufficient since differences in raw material procurement and manufacturing technique can lead to divergent breakage patterns for typologically identical vessel groups.

Estimated vessel equivalents (rim and base) – EVE

A number of different methods have been described in the bibliography to establish the number of vessel equivalents (as opposed to an estimate of vessels represented) in a given assemblage. For sites and periods with established average weights for specific vessel types (e.g., transport amphorae), the weight for different categories can be used to calculate the vessel equivalent (e.g. Baumhoff and Heizer 1959, 309; Raux 1998, 12). Similarly relying on standardized sizes, which are known to the researcher. are methods measuring surface area as a base for estimating vessel equivalents (e.g. Hulthén 1974; Byrd and Owens 1997). Other methods rely on the measurement of preserved rim, base, or even handle portions (Egloff 1973; Orton *et al.* 1993, 173); this last type of vessel-equivalent estimates is able to incorporate breakage rate data.

Outside of periods with highly standardized pottery production, the standardized weight EVE is of marginal value; intra- and inter-site variation of vessel dimensions, and consequently also weights, is high in the Late Bronze and Early Iron Age, in addition to a lack of assemblages with sufficient complete vessels of different shapes to establish a reliable average weight (Catling and Lemos 1990, 159 table 20, publish some vessel weights; their sample comprises 13 examples of eight shapes, only three of which are preserved whole).

The calculation of EVEs is here based on the measurement of the preserved portion (in %) of a vessel rim or base (for a description of the procedure, see Orton *et al.* 1993, 172–173; the method was first formulated by Egloff 1973). For calculation, the rim EVE, base EVE, or a combination of the two (in which case 720° equal one complete unit, rather than 360° for calculation of rim and base EVE separately) is used. 100% of rim, base, or rim+base circumference is taken to equal one vessel. Consequently, the figures obtained using this method are usually fractions, rather than whole numbers. It should be noted that the resulting data do not represent a reconstruction of the number of vessels represented, and that the greater the fragmentation and the lower the completeness, the smaller the EVE value in comparison with vessel numbers established by using MNIs.

At Kalapodi, the rim EVE was given preference over the base or rim+base EVE, because of a number of considerations. In the overall sample, rims outnumbered bases by almost 2:1 (Fig. 8) due to the greater breakage rate of rims in comparison to bases. Consequently, the representation of rims in the overall sample was much better than that of bases. In addition, while not every rim can be identified as to vessel shape, the rims in our sample were typologically more diagnostic than the bases. While adding rim and base measurements would increase the number of sherds included in the sample and thus, presumably, the accuracy of the data, we found this to be impractical at Kalapodi. A comparison of rim to base-ratios (Fig. 6) reveals that while for most vessel classes this is around 2:1, in the case of handmade cooking pots rims outnumbered bases by more than 5:1; similarly, the ratio of cooking pot handles to bases was about 4:1. Comparison of cooking pot rim breakage (as obtained through our EVE measurements, Fig. 7) indicates that this is not the result of exceptionally high breakage rates of cooking pot rims, which fall close to the average preservation rate of 6.7%, or 15 sherds per complete rim (handmade cooking pots: 6.5%). These findings strongly indicate that a majority of cooking pots at Kalapodi would have had round, rather than flat or otherwise articulated bases, which generally cannot be distinguished from body fragments among the sherd material. A calculation of vessel equivalents using the rim+base average would thus underestimate the number of cooking pots in the sample.

Therefore, the combined rim+base EVE is useful for units with little material, since the sampling base is increased; however, a correlation with other figures, such as the feature sherd ratio or the rim EVEs taken separately, is necessary to check for abnormalities.

The rim or base portion EVEs were criticized by Chase (1985), following his experimental study of the relationship between whole vessels and the sherds resulting from breakage. Specifically, he was unable to retrieve 100% of any vessel rim's circumference, and found that an accurate measurement of rim diameter on the basis of small sherds was not possible (Chase 1985, 217); Chase consequently dismissed the rim EVE method for estimating vessel numbers from sherd material. Note, however, that the average deviation for rim portions measured is 6°, or 1.7% (Chase 1985, table 1), which in most archaeological situations represent but a minor factor, since retrieval standards, sherd wear, and loss due to formation processes, all affect the degree of completeness of find material.

In our own sample, we observed a number of problems with the application of the rim EVE method similar to Chase's experience. For rim fragments preserving less than 5% of the rim circumference, the rim diameter could usually not be measured accurately, rendering the EVE value an educated guess rather than a true measurement. A comparison of rim diameter frequencies among open shapes showed a predominance of even-number diameters (10, 12, 14cm, etc.), which in all likelihood does not reflect the actual distribution of rim diameters in our sample, but rather results from the use of a diameter chart with even diameters in bolder print than uneven ones; this notion is supported by the fact that the pooling of even diameters is predominant for rims preserved below 5%.[5] Finally, despite our expectations of rim and base EVEs approximating one another in the overall sample, which at over 10,500 sherds we assumed to be large enough to constitute a representative sample, the figures for these two measures in the overall sample diverge by more than 25% (Fig. 8).[6] Examination of the raw data suggests that a number of possible explanations can be excluded, such as the over-representation of bases in our trench, repeated and major errors in measuring diameter and preservation, or regular under-estimation of rim portions for rim fragments under 5% preservation. While the latter factor might have contributed to the disparity in the figures, we consider it far more likely that the differential preservation of rims and bases is at the root of our figures. Thus, the average preserved base represents 18.6% of the total, compared to 6.6% for rims; in other words, bases break into an average of just over 5 sherds, compared to 15 sherds per rim. A significant number of rim sherds measured below 5% preservation (Fig. 9 shows this for small open shapes in our sample), and it can be expected that not all of these small and very small sherds were retrieved in excavation (Chase 1985, 215, used a half-inch screen to imitate retrieval in an archaeological situation). Moreover, according to our observations the rim tip (the outer 0.1–0.5cm) of a sherd is often lost, leading to a classification of the sherd as body, rather than rim sherd, while the resulting 'lip chip' is unlikely to have been recovered (as can be seen in the presence of near-rim body sherds and absence of 'lip chips' in our sample). These processes, rather innocuous and unavoidable in themselves, presumably are the main factor in the discrepancy between rim and base EVE. While not the case for the Kalapodi material, it can also be expected that wear at breaks would disproportionately affect the class with smaller sherd size (in this case, rims).

[5] Similarly, rim diameters measured in increments below 1cm occur from a minimum preserved rim portion of 4% in the present sample (0.5cm), even greater accuracy (to 0.2cm) above 7% preserved rim portion. Thus, accuracy of rim and base measurements cannot be expected below 1cm increments for highly fragmented assemblages. Note that in Kalapodi, handmade pottery represents a large part of the entire ceramic assemblage; for these wares, mostly cooking vessels and transport/storage jars, perfectly circular rims and bases are rare, and diameter measurements consequently often represent best guesses, even for large preserved portions. Similar to small sherds from wheelmade vessels, this shows that for fragmented assemblages diameters should be understood as trends, rather than accurate measurements of the ceramic material.

[6] Interestingly, the difference between rim and base EVE (55.00 and 72.67 for the entire trench) is disproportionately accounted for by phase IV (13.56 and 20.95, or a difference of 35%; Fig. 10), and particularly by small open vessels within phase IV (9.62 and 16.14, a difference of 40%). Why this is the case, is unclear at present. We expect that future research at Kalapodi, as well as at other sites studied by the author, will clarify the effects of breakage patterns, sherd size, and wear, as well as fabric and shape, on measurements of vessel rims and bases.

Minimum number of individuals – MNI

As for vessel equivalents, numerous methods for establishing the minimum number of individuals present in a given assemblage can be identified in the relevant literature. In an ideal situation, all sherds in an assemblage can be identified as belonging to one discrete vessel ('vessel batches', Baumhoff and Heizer 1959, 308, citing Newell and Krieger 1949; 'sherd families', Orton *et al.* 1993, 172); however, this approach relies upon a high degree of completeness, highly distinctive fabrics or other identifiable individualized features (e.g. decoration, see for example Eretria XX).[7] In less ideal situations, compromises have to found which can take the form of counts of diagnostic vessels parts (Raux 1998, 13 'nombre minimum pondéré'; Slane 2000, 378 '… in reality [MNI] is often simply a count of the number of different rims.'). However, it is often unclear whether the data presented as MNI do, in fact, represent a minimum or a maximum number of individuals, or a number somewhere in between, particularly for sherd assemblages.

Due to the characteristics of the Kalapodi assemblage, with low completeness, little variation in decoration and gradual differences in fabric, the sorting of sherds into 'families' was not practicable. Consequently, we attempted to identify diagnostic vessel parts that could be equated with one vessel, and that ideally could be subjected to typological classification (for a similar approach, see also van de Moortel 2001, 29 n. 15); the results of this effort are listed in Fig. 12. As becomes obvious from Fig. 12, in a fragmentary state most shapes are best identified by their handles. Rim types are shared between cups and skyphoi (and, depending on the size of the preserved fragment, even angular bowls and possibly kylikes), and amphorae, hydriae, and jugs; the same is the case for base types. Breakage patterns observed at Kalapodi and elsewhere suggest that the attachment of the handle to the body is one of the strongest point in a vessel, unless it had been made very carelessly; the handle itself often breaks into a varying number of sections, for which reason attachments only, rather than any handle fragment, are being considered. To obtain a minimum number of vessel represented, the number of diagnostic pieces is then calculated as illustrated in the following example.

Wheelmade painted small open shapes:
- vertical handles complete (1)
- fragments of vertical handles attached at rim (5)
- fragments of vertical handles attached to body (2)
- horizontal handles complete (2)
- fragments of horizontal handles attached to body (6).

→ Total vertical handles: 6 (1+5; among the attachments to rim and body, the higher number is used for calculation).
→ Total horizontal handles: 5 (2+[6/2]; joins between the attached fragments cannot be ruled out, thus for this calculation two attachments equal one complete handle).

[7] It should be noted that the latter methods disadvantage undecorated vessel classes, such as cooking and utilitarian wares, thus rendering them impractical for the analysis of entire assemblages that include coarse as well as finewares.

→ MNcups: 6 (1 vertical handle each), MNskyphoi: 3 (two horizontal handles each); since the MNI counts operate without fractions, the 1 spare horizontal handle does represent 1, rather than 0.5, vessel.
→ In addition, a MaximumNI can be calculated by equating each of the handle parts with one vessel (in this example, MaxNcups: 8, MaxNskyphoi: 8).

While adding somewhat to the sorting and recording process (as handle attachments, complete handles, and handle segments have to be counted separately), this system allows for fairly accurate recording of vessel individuals, based on formal criteria that are easily defined and replicated. Especially for calculating numbers of different small open shapes, we found this method useful.

The problems this approach poses, however, are equally clear from Fig. 12. Among large closed shapes, there is considerable overlap between handle types (amphora, hydria, and jug can all have vertical handles). At the current stage of research, the size ranges for large closed shapes are not yet clear; consequently, a sorting of handles into 'small', 'medium', and 'large' handles, corresponding to different shapes (e.g. 'small' vertical handle for jug, 'large' vertical handle for neck-handled amphora), has to wait for a better definition of the types represented at Kalapodi.

Aggregate feature count

The aggregate feature count consists of the added sums for rim, handle, and base sherds in each ceramic category. It was found that the percentages obtained through the AFC behave in ways similar to the EVE figures, rather than the simple sherd count or weight (Figs. 11 and 13). While not resulting in figures which can be used to calculate the number of individuals present in a given deposit, the AFC is a simple and time-saving means to obtain percentages for ceramic distribution more closely reflective of the actual distribution of vessels in a sample than count or weight.

The assumption underlying this suggestion is that irrespective of size and type, given a similar depositional history, different vessel classes break into comparable numbers of feature sherds (which is here taken to denote structural features, that is, rims, handles, and bases). In addition, it is posited that virtually all vessel classes present actually have rims, handles, and bases; in a hypothetical assemblage of handleless bowls and jars with four handles, the feature count would obviously be skewed in favour of the latter. The first assumption, i.e. similar depositional history, should by definition be fulfilled within a deposit or SU (though see Schiffer 1987, 265–267). The second assumption appears to be justified empirically – the vast majority of shapes commonly identified in Bronze and Iron Age assemblages come with identifiable bases, handles and rims. No independent means of testing the validity of the AFC is available, since our sample comes from a population of unknown size and composition, and no experimental data, e.g. from smashing a set of complete pots in order to observe breakage patterns, was available. However, using the figures obtained by measuring vessel rims and bases, it is suggested that in the Kalapodi sample breakage results in about 32 feature sherds per vessel, irrespective of vessel type (Fig. 14).[8] There are significant deviations from this average; most of these, however, occur in poorly represented categories and can be explained as sampling error, rather than a meaningful variation. Handmade cooking pots are an exception (49.2 features per vessel); these account for 10.1% (rim EVE) of the total ceramics in the sample, and thus are among the more common classes. Comparison of rim preservation for cooking pots and the overall average (the average cooking pot rim represents 6.6% of the total circumference, the same as the overall average, Fig. 7) suggests that the explanation for this deviation is not to be sought in a divergent breakage pattern. In fact, the number of features per vessel comes close to the overall average when using the rim EVE (29.4 feature per pot), rather than the average from rims and bases.

Thus, reasons for the divergence of figures for cooking pots have to be sought elsewhere: together with the significantly different rim-to-base ratio for cooking pots (Fig. 6), the features-per-pot figures demonstrate that the majority, if not all, of cooking pots at the site had rounded bases; sherds from these bases generally cannot be distinguished from body sherds and thus lead to under-representation of cooking pot bases when compared to other categories.

COMPARISON OF RESULTING DATA SETS

Sherd counts and weights currently account for the vast majority of quantitative studies available for the Aegean, and some scholars have suggested that these figures are adequate and sufficient for the characterization of assemblages (e.g. Tomber 1993, 155; Slane 2003, 321 n. 2).[9] That sherd counts and weights can, however, be problematic when used for estimating the actual representation of vessel classes is clearly illustrated by an example from the Kalapodi assemblage.

In phases III and IV, handmade coarseware accounts for 24.1% and 33.0% of the sherd count, 26.7% and 44.6% of weight respectively (Fig. 15). Thus, sherd count and weight appear to indicate a substantial increase of handmade coarsewares in the Middle to Late Geometric phase. However, both AFC and rim EVE put this category at a much lower frequency, with 13.7% and 15.4% of AFC,

[8] The figure was obtained by dividing the total feature count for each ceramic category by its average EVE (rim + base EVE/2).

[9] Tomber states that percentages based on counts of rims, handles, and bases perform very similar to those based on rims, handles, bases, and body sherds, suggesting that the two types of count are essentially equivalent, and that consequently deposits with figures based on total count and AFC (in our terminology) can be compared. Note, however, that Tomber's figures are based on a subset of classes within assemblages, namely amphorae. In our experience, assemblages with mixed composition (fine, coarse, cooking wares; range of sizes) do behave rather differently; a glance at the numbers of feature sherds per vessel across different ceramic classes (Fig. 3), as well as the wide range of total sherds per vessel (Fig. 5) immediately show the problem of comparing Total count and AFC for most deposits (see also Fig. 17–18 graph B).

or 7.9% and 16.0% of EVE values for phases III and IV respectively (Fig. 16). Similarly, handmade cooking pots appear to increase significantly in sherd count and weight from phase II to phase III, while both AFC and rim EVE indicate a decline in actual numbers. Thus, handmade coarse and cooking wares, rather than increasing dramatically in the 9th and 8th centuries BC (phase IV), appear to have decreased in comparison with phase II. The difference between frequencies of sherds of handmade wares in phase II on the one hand, and phases III and IV on the other, is consequently better explained by an increase in individual vessel size and weight, rather than in actual vessel numbers.

Consequently, statements regarding the dramatic increase of handmade wares in Early Iron Age levels at Kalapodi, and particularly that handmade wares comprise more than half of the pottery found (Jacob-Felsch 1996, 73; Lemos 2002, 86), have to be used with caution as they derive from sherd counts only and do not take into account the representation of vessel individuals.[10]

Despite the criticism of the rim and base EVE method voiced by Chase (1985) and others, in our experience this method was the most useful out of the set tested at Kalapodi. The 'counting' of sherds by increment, rather than equating a random and variably preserved portion of a vessels with '1' (as done for MNI counts, see below; Orton 1982, 1), allows for the incorporation of degrees of breakage, and thus greatly increases accuracy, since differential breakage patterns between vessel classes, when using the rim/base EVE method, do not affect the resulting data (similar conclusion by Orton 1982, 14). The application of this method allows for the processing of samples of any size and composition (i.e., individual lot numbers, as well as whole SUs), since the resulting data can simply be added up for each SU (or other larger unit, e.g. our assessment of the material from the entire trench in Fig. 11); corrections and changes to the interpretation of the site stratigraphy consequently do not necessitate a partial or complete re-processing of the ceramic material. Particularly for sites with limitations on space and access, this greatly enhances the applicability of the rim/base EVE method.

Our expectation was to find that the various methods estimating vessel numbers, as opposed to quantity of sherds, would yield roughly comparable results, and that differences would decrease as sample size increased. That the correlation between data sets resulting from different quantitative methods is not entirely straightforward can already be seen in the differences for rim and base EVEs from the same SU (*supra*; Figs. 8 and 10). In order to investigate the correlation between methods beyond the EVEs, a comparison was made between rim and base EVE, AFC, and the rim count (as a basic means of determining the MNI; Slane 2000, 378). The standard deviation between the percentages for the four data sets was calculated and the results ranked (Fig. 18).[11] Initially, the results appeared promising since the standard deviation for more than 60% of values falls below 2%. However, when comparing standard deviation and sample size (here the total sherd count underlying each value) the correlation between the two was not as expected (Fig. 18, graph C); the high number of values below 2% deviation derived from ceramic categories with very little material, i.e., categories that in any counting method would be sparsely represented. Conversely, large numbers of sherds did not ensure comparable values for EVEs, etc., since even for categories containing in excess of 700 sherds, deviation could be more than 10%. Since the differences between rim and base EVEs were already clear, the same calculation was performed again, this time however excluding the base EVEs. The resulting figures show that while the concentration of low values for the standard deviation for small categories persists, all categories now show deviations below 5% (Fig. 18, graph D).[12] Handmade cooking pots, which for the previous calculation provided several of the highest values in deviation, now fall within the overall average of values. This reinforces our notion that the majority of handmade cooking vessel at Kalapodi were made with round bases; since cooking pots account for about a quarter of vessels in phases II and III (Fig. 16), the 'missing' cooking pot bases distort the data for the entire assemblage when the base EVEs are taken into account.

These findings would suggest that comparison between figures with different underlying methodologies (from the range of methods described for establishing vessel equivalents and MNIs), while not entirely unproblematic, should be viable particularly in studying broad trends across assemblages; however, the collection of more data is desirable to observe how the different methods perform in relation to one another in different assemblages.

Since no independent means are available to test the accuracy of any of the methods here assessed (though experiments with breakage and retrieval, along the lines of Chase's work (1985), might prove of value), no conclusion can be drawn as to which is the 'right' method. Each has benefits and drawbacks, and will be chosen depending on the individual researcher's preference, while to no small degree being influenced by the characteristics of the assemblage under study. As shown above, dependence on sherd counts and weights alone is insufficient for the study of ceramic use patterns and their development, and the

[10] Neither Jacob-Felsch nor Lemos claim that the figures discussed represent numbers of vessels, rather than numbers of sherds; in fact, this issue is not broached at all by either scholar. In the absence of commentary, however, the reader's intuitive response is to understand these figures as reflecting actual representation, that is, vessel numbers. However, as shown above, both the representation of vessel classes, and diachronic development of representation, can be grossly misleading when based solely on sherd count and/or weight.

[11] The standard deviation is calculated using the function included in Open Office's spreadsheet package. The data compared here are percentages for each ceramic category in phases I-IV, excluding categories without sherds (e.g. the category 'wheelmade plain ware' is absent in phases II-IV); for the methods compared in graphs B-D in Fig. 18, see the individual labels.

[12] A comparison with data collected by the author from roughly contemporary wells in the Athenian Agora shows that this degree of deviation is not limited to Kalapodi; at the Agora, only AFC and rim count are available to date, and even though the rim count is essentially a subset of the AFC, standard deviation in categories containing more than 700 sherds was still close to 4%.

employment of a method calculating vessel numbers, and thus the actual numerical relation between vessel classes, already represents a great improvement. Here again, transparency and diligence in the presentation of data in publication might go a long way to ensure the usefulness of such data in the wider context of research.

THE RELATION BETWEEN SAMPLE SIZE, DISTRIBUTION, AND REPRESENTATION OF ARTEFACT CLASSES

One major consideration when applying quantitative methods to the analysis of archaeological data is the degree to which the material studied, which usually is merely a sample of a larger whole, represents its parent population; the validity of conclusions drawn from our analyses, as well as the formulation of research questions based on this data, is built upon the reliability, that is the accuracy and validity, of the data obtained (see Orton 1982, 6, for some interesting thoughts regarding excavation as a type of sampling). A determination of what portion of the parent population is excavated is often not possible (O'Neil 1993 suggests that even large percentages of a known population are not necessarily representative). The extent of a site, or of an excavated deposit within a site, is not always known well enough to calculate the ratio between parent population and sample; even closed deposits, such as pits, wells, or burials, can be understood as samples of a larger parent population, for example of all pottery from the same site, region, or chronological phase, thus raising the question whether quantitative data from these contexts have wider applicability beyond a merely descriptive summary of a deposit's contents.

Sample size, however, also denotes the actual size of a sample at hand, and it is this meaning of the phrase that is more pertinent to the discussion here; a number of stipulations can be made at this level. Of course, the larger the sample the greater the likelihood that the distribution of categories within that sample resembles the parent population; but how to determine an adequate sample size? Madrigal suggests that no category within a sample should contain less than five individuals (Madrigal 1998, 193; cf. Slane 2003, 324, citing Riley 1979 who considers 60–100 feature sherds as the minimum size for a deposit to be considered statistically significant; Tomber 1993, 155, uses a minimum 260 sherds in her discriminant analysis). Quite clearly the number of categories identified, and the evenness of their distribution, have a very noticeable effect on the sample size required to ensure adequate representation of all categories; thus Orton (1982, 17–18) suggests an EVE value of 1 for the smallest pottery category recognized as the minimum sample size.
An example illustrating the effect of number of categories and sample size on distribution is here presented based on the Kalapodi finds. For the Early Iron Age, nine ceramic categories can be defined according to our sorting system (wheelmade painted small and large open, small, medium and large closed, handmade coarse, handmade cooking, pithos, other). If these nine categories were evenly distributed, about 50–100 sherds would be needed in any sample simply to ensure that all categories were represented, let alone in their 'true' numerical proportion.[13] In reality, categories are rarely distributed evenly, thereby increasing the adequate sample size considerably. In the Kalapodi example, only six lots (out of 44) contain examples of all nine categories (Fig. 19). The rarest group here are wheelmade painted small closed vessels, which in phases II-IV comprise 0.1/0.2/1.2% respectively. The six lots containing examples of all nine ceramic categories counted to a total of 98, 258, 298, 473, 482, and 499 sherds respectively. However, while two lots with less than 300 sherds did contain examples of all categories, another five lots in the same size range contained only seven or eight categories. It is only with lot sizes close to 500 sherds that these nine ceramic classes are typically all present. When incorporating different sherd types (rim, handle, base, body), 36 sorting possibilities result (Fig. 20). The three largest lots now score between 21 and 26 out of a total 36 categories possible. When the sherd counts for the three Iron Age phases are added up, these larger groups (1098, 2033, 2599 sherds) score between 30 and 33. Only all Iron Age phases added together, for a total of 5730 sherds, contain examples of all 36 categories defined for this example. A glance at the absolute counts shows that, in fact, a number of categories are represented by single sherds only (Fig. 21).

Looking at the issue from a different angle, the smallest category (wheelmade painted small closed) is represented with 0.5% of the count, thus on average, one in 200 sherds is identified as this type. 19.4% (or 0.1% of the total) of these are feature sherds – that is one small closed feature in 1000 sherds. The probabilities for specific features are even smaller; thus, among 5730 sherds, only one small closed handle fragment was identified. Thus, based on the known distribution of vessel categories at a site, predictions can be made as to the minimum sample size needed to make reasonably safe statements regarding rare vessel classes; the less even the distribution of classes, the larger the sample needs to be to ensure representation.

Two conclusions can be drawn from this – first, adequate sample size is larger than we would have expected at Kalapodi, and given the rarity of some classes, we still cannot ensure that the distribution of ceramics calculated for our sample trench accurately reflects the population; second, for samples with uneven distribution, or small sample sizes, a multitude of sorting categories leads to unreliable results since accurate representation cannot be guaranteed. While it is not clear that this can be adequately addressed, for example by choosing broader sorting/recording categories as these are dependent on the aims of the study, it should be kept in mind that rare object classes are affected disproportionately by sample size.

[13] An attempt to calculate the minimum number of sherds needed to identify all nine categories had the following result: for 95% probability, a sample size of 64 is required, while 91 sherds are required for 99% probability ($n = \ln(1-p)/\ln(8/9) + \ldots + \ln(1-p)/\ln(1/9)$). Many thanks are due to Mario Strack and Herbert Strack for advice on the calculation of these figures.

SUMMARY OF RESULTS

The application of a range of quantitative methods, including sherd count, weight, and methods establishing vessel equivalents and minimum numbers of vessels represented, gave us the opportunity to assess the usefulness of different means of recording and analyzing sherd assemblages. The results of our trials are summarized in the following.

Sherd counting and weighing are useful and speedy methods of recording and provide a basis for a rough initial estimate of the representation of vessel classes in a given deposit. Particularly when time is short, weighing of sherds is a valuable method for obtaining accurate quantitative data. However, for analysis of assemblages with a view toward vessel function, as well as for diachronic and inter-site comparison, too many biasing factors in simple weights and counts render these methods unreliable and are too complex to control for. In fact, our analysis has shown that sherd counts and weights can result in poor data when characteristics of a typological vessel class change over time. Furthermore, it is posited that while many published quantitative analyses of ceramic groups make use of mere sherd counts and/or weights, the resulting figures are often intuitively understood to represent not the amount of pottery, but the true numerical relation between shapes. In these circumstances, methods aiming at establishing vessel numbers, rather than sherd quantities, yield more accurate and reliable results.

Particularly for highly fragmentary assemblages, vessel equivalents (EVEs) based on measurements of preserved portions of rims and bases were found to be very useful. First, measurement of preserved rim portions takes into account breakage rates and thus allows a control for varying degrees of breakage between different assemblages, but also between different vessel classes within one assemblage. Even though the latter does not appear to apply at Kalapodi, since figures for average preserved rim portions are relatively even across vessel categories, this feature of EVE calculations is very important for assemblages with mixed contents comprising fineware, coarseware, and cooking wares. Second, EVE values can be obtained for random sections of a SU and simply added up; particularly at sites with complex stratigraphies, and non-closed deposits not excavated in their entirety, this allows for continued study of excavation records etc. for the purpose of reconstruction of stratigraphy outside precious apotheke-time.

Calculation of MNIs based on the 'sherd family' approach (Orton *et al.* 1993, 172) was found to be too onerous for the degree of fragmentation present in the Late Bronze and Early Iron Age levels at Kalapodi, and was consequently not attempted. Rim counts, which are proposed as a solution for fragmentary assemblages by Slane (2000, 378) performed well in comparison with other methods, but are rather susceptible to problems caused by differential breakage patterns between different vessel classes.

In addition, in many instances in the bibliography MNIs do not, in fact, represent the minimum number of individuals identified, nor even the maximum number, but rather an unidentifiable figure in between. This renders the resulting figures inherently unreliable; the 'bonus' of being able to investigate actual numbers of vessels represented does not off-set this problem, and in fact could lead to substantial over-estimation of vessel numbers. Moreover, while minimum represented numbers allow for a more intuitive understanding of vessel quantities, in many instances these figures are, in fact, meaningless, when they refer to samples from an unknown, larger population, deriving for example from fills and dumps.

While clearly yielding more accurate data sets, methods used for establishing EVEs and MNIs are, in general, fairly time-consuming, and particularly in the case of MNI counts based on the identification of 'sherd families' among fragmented assemblages, may require a good deal of initial training. We have attempted to show that the significantly less labour-intensive aggregate feature count performs similar to EVE and MNI values and thus allows for assessment of vessel representation.

The comparison between different methods of quantification has shown that while differences persist between datasets derived from the same assemblage, the overall trends in assemblage composition are reflected by all methods here assessed. This result makes us hopeful that even without a common methodology, assemblages from different sites and periods can be studied in a comparative framework as long as the distinction between quantities of sherds, and quantities of vessels, is understood and made explicit in published data.

Finally, the use of different methods of quantification and comparison of results for different ceramic categories and phases has indicated that these methods can be very useful for understanding characteristics of the ceramic assemblage not immediately apparent when studying the material. Thus, the presence of round-based cooking pots at Kalapodi was suspected, due to the lack of cooking pot bases among the finds studied so far, but could hardly be demonstrated without complete cooking pot profiles. Utilizing both ratio of feature sherds and base EVEs, we could conclusively prove the existence of this type at Kalapodi.

At the same time, this example shows the value of using more than one quantitative approach, at least on an initial good-sized sample, in order to evaluate the best-suited approach, and to assess the significance of irregularities that emerge in quantitative datasets. At Kalapodi, the preparations for the 2009 'Quantitative approaches' workshop presented us with the opportunity, at times in a somewhat playful manner and by means of trial and error, to observe the performance of different quantitative methods for this particular assemblage, and provided a very clear basis for the formulation of future research policies and methodology at the site.

Ackowledgements

This paper resulted from discussions during and after the workshop 'Early Iron Age pottery: a quantitative approach' organized by S. Verdan, T. Theurillat, and A. Kenzelmann Pfyffer and I would like to thank the organizers for agreeing to publish this paper in the proceedings of said workshop. Furthermore, many thanks are due to Ivonne Kaiser, Benjamin Millis, and Jeremy Rutter for reading, commenting on, and improving earlier versions of this text; all errors and omissions remain solely my own responsibility.

References

Arcelin, P. and Tuffreau-Libre, M. (eds), 1998. La quantification des céramiques: conditions et protocole. Actes de la table ronde (Centre archéologique européen du Mont Beuvray, 1998), Glux-en-Glenne.

Baumhoff, M.A. and Heizer, R.F., 1959. Some unexploited possibilities in ceramic analysis, Southwestern Journal of Anthropology 15, 308–316.

Bobrowsky, P.T. and Ball, B.F., 1989. The theory and mechanics of ecological diversity in archaeology, in Leonard and Jones 1989, 4–12.

Byrd, J.E. and Owens Jr., D.D., 1997. A method for measuring relative abundance of fragmented archaeological ceramics, JFA 24, 315–320.

Cannon, A. 1985. The Quantification of Artifactual Assemblages: Some Implications for Behavioral Inferences, AmerAnt 48, 785–792.

Catling, R.W.V. and Lemos, I.S., 1990. Lefkandi II.1: The Protogeometric Building at Toumba, London.

Chase, P.G., 1985. Whole vessels and sherds: an experimental investigation of their quantitative relationships, JFA 12, 213–218.

Egloff, B.J., 1973. A method for counting ceramic rim sherds, AmerAnt 38, 351–353.

Eretria XX = S. Verdan, A. Kenzelmann Pfyffer, C. Léderrey, Céramique géométrique d'Erétrie (Eretria. Fouilles et recherches, vol. XX), Gollion 2008.

Felsch, R.C.S., 1987. Kalapodi. Bericht über die Grabungen im Heiligtum der Artemis Eplaphebolos und des Apollon von Hyampolis 1978–1982, AA, 1–26.

Gifford, E.W., 1951. Archaeological excavations in Fiji, Anthropological Records 13, 189–288.

Hulthén, B., 1974. On choice of element for determination of quantity of pottery, Norwegian Archaeological Review 7, 1–5.

Jacob-Felsch, M. 1987. Kalapodi Bericht 1978–1982. Bericht zur spätmykenischen und submykenischen Keramik, AA, 26–35.

Jacob-Felsch, M., 1996. Die spätmykenische bis frühprotogeometrische Keramik, in R.C.S. Felsch (ed), Kalapodi. Ergebnisse der Ausgrabungen im Heiligtum der Artemis und des Apollon von Hyampolis in der antiken Phokis. Band I, Mainz, 1–213.

Jones, G.T. and Leonard, R.D., 1989. The concept of diversity: an introduction, in Leonard and Jones 1989, 1–3.

Kintigh, K.W., 1984. Measuring archaeological diversity by comparison with simulated assemblages, AmerAnt 49, 44–54.

Kintigh, K.W., 1989. Sample size, significance, and measures of diversity, in Leonard and Jones 1989, 25–36.

Knappett, C., 2001. Overseen or Overlooked? Ceramic Production in a Mycenaean Palatial System, in S. Voutsaki and J. Killen, Economy and Politics in the Mycenaean Palaces States. Proceedings of a Conference held on 1–3 July 1999 in the Faculty of Classics, Cambridge (Cambridge Philological Society Supplementary Volumes no. 27), Cambridge, 80–97.

Lemos, I.S., 2002. The Protogeometric Aegean. The archaeology of the late eleventh and tenth centuries BC, Oxford.

Leonard, R.D., 1997. The sample-size richness relation: A comment on Plog and Hegmon, AmerAnt 62, 713–716.

Leonard, R.D. and Jones, G.T. (eds), 1989. Quantifying diversity in archaeology, Cambridge.

Madrigal, L., 1998. Statistics for anthropology, Cambridge.

Meltzer, D.J., Leonard, R.D. and Stratton, S.K., 1992. The relationship between sample size and diversity in archaeological assemblages, JAS 19, 375–387.

Newell, H.P. and Krieger, A.D.,1949. The George C. Davis Site, Cherokee County, Texas, Menasha/Wisconsin.

Nitsche, A., 1987. Protogeometrische und Subprotogeometrische Keramik aus dem Heiligtum bei Kalapodi, AA, 35–49.

O'Neil, D.H., 1993. Excavation sample size: A cautionary tale, AmerAnt 58, 523–529.

Orton, C., 1982. Computer simulation experiments to assess the performance of measures of quantity of pottery, World Archaeology 14, 1–20.

Orton, C., 1993. How many pots make five? – An historical review of pottery quantification, Archaeometry 35, 169–184.

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

Raux, S., 1998. Méthodes de quantification du mobilier céramique. Etat de la question et pistes de réflexion, in Arcelin and Tuffreau-Libre 1998, 11–16.

Riley, J.A., 1979. The coarse pottery from Berenice, in J.A. Lloyd (ed), Excavations at Sidi Khrebish Benghazi (Berenice) II, Tripoli, 91–467.

Rutter, J.B., 1990. Pottery groups from Tsoungiza of the end of the Middle Bronze Age, Hesperia 59, 375–458.

Schiffer, M.B., 1987. Formation processes of the archaeological record, Salt Lake City.

Shaw, J.W., Van de Moortel, A., Day, P.M. and Kilikoglou, V., 2001. A LM IA Ceramic Kiln in South-Central Crete: Function and Pottery Production (Hesperia Supplement 30), Princeton.

Slane, K.W., 2000. Review of P. Arcelin and M. Tuffreau-Libre (eds), La quantification des céramiques: Conditions et protocole, AJA 104, 377–378.

Slane, K.W., 2003. Corinth's Roman pottery: quantification and meaning, in C.K. Williams II and N. Bookidis (eds), Corinth: The Centenary 1896–1996 (Corinth XX), Princeton, 321–335.

Solheim, W.G. II, 1960. The use of sherd weights and counts in the handling of archaeological data, Current Anthropology 1, 325–329.

THOMAS, P.M., 2005. A Deposit of Late Helladic IIIB1 Pottery from Tsoungiza, Hesperia 74, 451–573.

TOMBER, R., 1993. Quantitative approaches to the investigation of long-distance exchange, JRA 6, 142–166.

VAN DE MOORTEL, A., 2001. The Area around the Kiln, and the Pottery from the Kiln and the Kiln Dump, in J.W. Shaw *et al.*, 25–110.

WELLS, B. 1983. Asine II. Results of the excavations east of the Acropolis 1970–1974. Fasc. 4: The Protogeometric Period (Part 2: An analysis of the settlement; Part 3: Catalogue of pottery and other artifacts; Part 2 and 3 numbered consecutively), Stockholm.

Abbreviations used in the figures

WM wheelmade
HM handmade
ptd painted
pl plain
s/ m/ l/ small, medium, large
/o /c open, closed
R, H, F, B rim, handle, foot (base), body
Ct, Wt count, weight
EVE estimated vessel equivalent
AFC aggregate feature count (rims+bases+handles)

Fine		**wheelmade**	**painted**	**open**	**small**
					large
				closed	**small**
					large
			plain	**open**	**small**
					large
				closed	small
					large
		handmade			
Coarse	**Utilitarian**	**wheelmade**	**painted**	open	small
					large
				closed	small
					large
			plain	open	small
					large
				closed	small
					large
		handmade	painted		
			plain	**open**	**small**
					large
				closed	small
					large
	Cooking	**wheelmade**			
		handmade	**local/regional**		
			imported		
	Pithos	wheelmade			
		handmade	**decorated**		
			plain		

Fig. 1. Kalapodi sorting concept. Types in bold face have been identified at the site.

Fig. 2. Kalapodi trench 5030/4965, all lots (LHIIIC-LG)	R (Ct)	H (Ct)	F (Ct)	B (Ct)	total Ct	total Wt
WM ptd s/o	453	305	281	1966	3005	12.237
WM ptd l/o	16	7	22	146	191	2.345
WM ptd s/c	5	6	5	76	92	0.284
WM ptd m/c	16	23	16	292	347	1.754
WM ptd l/c	21	18	20	731	790	7.416
WM pl s/o	120	34	45	379	578	1.416
WM pl l/c	4	4	2	90	100	0.403
WM coarse	4	1	4	101	110	1.337
HM coarse	67	74	59	1720	1920	27.519
WM cooking	35	10	19	554	618	4.176
HM cooking	83	64	16	1304	1467	16.455
pithos (-ware)	14	1	12	871	898	52.719
other	12	2	3	393	410	5.097
TOTAL	850	549	504	8623	10526	133.158

Fig. 2. Kalapodi trench 5030/4965, excavation seasons 2005–6, overall sherd count and weig ht for trench. Figures exclude six pre-sorted lots.

Fig. 3. Ratio of feature sherds to total sherd count	
WM ptd s/o	1.9
WM ptd l/o	3.2
WM ptd s/c	4.8
WM ptd m/c	5.3
WM ptd l/c	12.4
WM pl s/o	1.9
WM pl l/c	9.0
WM coarse	11.2
HM coarse	8.6
WM cooking	8.7
HM cooking	8.0
pithos (-ware)	32.3
other	23.1

Fig. 4. Average sherd Wt [g]	
WM painted small open	4.1
WM painted large open	12.3
WM painted small closed	3.1
WM painted medium closed	5.1
WM painted large closed	9.4
WM plain small open	2.4
WM plain large closed	4.0
WM coarse	12.2
HM coarse	14.3
WM cooking	6.8
HM cooking	11.2
pithos (-ware)	58.7
other	12.4
overall average	12.7

Fig. 5. Number of sherds per vessel unit	A	B	C
WM ptd s/o	101	100	1.0%
WM ptd l/o	398	180	0.6%
WM ptd s/c	161	281	0.4%
WM ptd m/c	347	329	0.3%
WM ptd l/c	401	561	0.2%
WM pl s/o	76	73	1.4%
WM pl l/c	476	379	0.3%
WM coarse	786	421	0.2%
HM coarse	389	433	0.2%
WM cooking	322	266	0.4%
HM cooking	264	266	0.4%
pithos	1575	947	0.1%

Fig. 6. Feature sherd ratios	rim	handle	foot
WM ptd s/o	1.6	1.1	1
WM ptd l/o	0.7	0.3	1
WM ptd s/c	1.0	1.2	1
WM ptd m/c	1.0	1.4	1
WM ptd l/c	1.1	0.9	1
WM pl s/o	2.7	0.8	1
WM pl l/c	2.0	2.0	1
WM coarse	1.0	0.3	1
HM coarse	1.1	1.3	1
WM cooking	1.8	0.5	1
HM cooking	5.2	4.0	1
pithos	1.2	0.1	1
overall	1.7	1.1	1

Fig. 3. Number of feature sherds (rims, handles, bases) in comparison to body sherd count for each ceramic category (number of features set as baseline).
Fig. 4. Comparison of average sherd weight (in gram) as indicator of brokenness for the ceramic categories identified at Kalapodi; average values for phases I-IV.
Fig. 5. Reconstruction of the number of sherds per vessel unit. A: total count/rim EVE; B: calculated with the overall average rim preservation (6.6%) and the percentage of rims of the total count for each category; C: percentage of the total vessel represented by an average single sherd (for figure in column B).
Fig. 6. Feature sherd ratio, comparing number of rims, handles, and bases. Number of foot/base fragments set as baseline (=1).

Fig. 7. Trench 5030/4965, rim EVEs				
category	rim sum	% sum	av. %	rim EVE
WM painted small open	461	2985%	6.5%	29.85
WM painted large open	13	48%	3.7%	0.48
WM painted small closed	4	57%	14.3%	0.57
WM painted medium closed	10	100%	10.0%	1.00
WM painted large closed	22	197%	9.0%	1.97
WM plain small open	121	766%	6.3%	7.66
WM plain large closed	4	21%	5.3%	0.21
WM coarse	3	14%	4.7%	0.14
HM coarse	62	494%	8.0%	4.94
WM cooking	31	192%	6.2%	1.92
HM cooking	84	555%	6.6%	5.55
pithos (-ware)	12	57%	4.8%	0.57
other	6	14%	2.3%	0.14
TOTAL	833	5500%	6.6%	55.00

Fig. 8. Rim and base EVEs	rim EVE	base EVE	rim+base EVE
WM ptd s/o	29.85	42.70	36.28
WM ptd l/o	0.48	1.30	0.89
WM ptd s/c	0.57	3.64	2.11
WM ptd m/c	1.00	2.16	1.58
WM ptd l/c	1.97	2.38	2.18
WM pl s/o	7.66	4.75	6.21
WM pl l/c	0.21	0.17	0.19
WM coarse	0.14	0.80	0.47
HM coarse	4.94	8.16	6.55
WM cooking	1.92	3.70	2.81
HM cooking	5.55	1.08	3.32
pithos (-ware)	0.57	1.66	1.12
other	0.14	0.17	0.16
TOTAL	55.00	72.67	63.84

Fig. 7. Rim EVEs for entire trench.
Fig. 8. Rim and base EVEs; average of rim and base EVE [(rim+base)/2]. Figures for entire trench (LHIIIC-LG), excluding six pre-sorted lots.

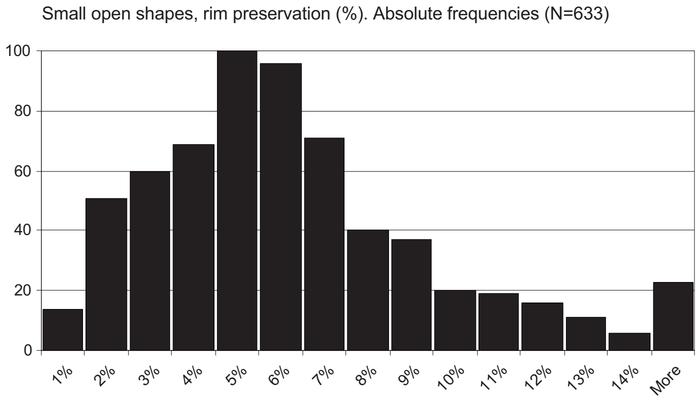

Fig. 9. Kalapodi trench 5030/4965. Rim preservation of small open shapes.

Phase	rim EVE	base EVE
I	18.83	22.45
II	2.79	4.97
III	9.54	8.62
IV	13.56	20.95

Fig. 10. Overall rim and base EVE figures for phases I-IV.

Representation of ceramic categories								
	total Ct	total Wt	total FeatCt	rim EVE	%Ct	%Wt	%AFC	% rim EVE
WM ptd s/o	3005	12.237	1039	29.85	28.5%	9.2%	54.6%	54.3%
WM ptd l/o	191	2.345	45	0.48	1.8%	1.8%	2.4%	0.9%
WM ptd s/c	92	0.284	16	0.57	0.9%	0.2%	0.8%	1.0%
WM ptd m/c	347	1.754	55	1.00	3.3%	1.3%	2.9%	1.8%
WM ptd l/c	790	7.416	59	1.97	7.5%	5.6%	3.1%	3.6%
WM pl s/o	578	1.416	199	7.66	5.5%	1.1%	10.5%	13.9%
WM pl l/c	100	0.403	10	0.21	1.0%	0.3%	0.5%	0.4%
WM coarse	110	1.337	9	0.14	1.0%	1.0%	0.5%	0.3%
HM coarse	1920	27.519	200	4.94	18.2%	20.7%	10.5%	9.0%
WM cooking	618	4.176	64	1.92	5.9%	3.1%	3.4%	3.5%
HM cooking	1467	16.455	163	5.55	13.9%	12.4%	8.6%	10.1%
pithos (-ware)	898	52.719	27	0.57	8.5%	39.6%	1.4%	1.0%
other	410	5.097	17	0.14	3.9%	3.8%	0.9%	0.3%
TOTAL	10526	133.158	1903	55.00				

Fig. 11. Representation of ceramic categories, listing absolute figures and percentages for each quantitative method.

Shape	**Identification from sherds**	**Diagnostic vessel part**
cup	small open shape with one vertical handle to rim	vertical handle, attached to body or rim
shallow angular bowl	small open shape with horizontal band handle below rim, carinated body	horizontal band handle, attached to body
kylix	small open shape with tall stemmed foot, variety of rim and handle types	lower part of bowl with stem attachment
skyphos	small open shape with two horizontal roll handles	roll handle, attached to body (2 complete or 4 attachments to one vessel)
krater	large open shape with horizontal roll handles on body	horizontal roll handle, attached to body
amphora	medium to large closed shape with two vertical handles to rim or neck	large vertical handle, attached to body, rim, or neck
jug	medium closed shape with one vertical handle to rim	medium vertical handle (band handle), attached to body or rim
trefoil jug	medium closed shape with one handle to rim, lip pinched into trefoil shape	medium vertical handle (band handle), attached to body or rim, lip near handle with distortion from pinching
hydria	medium or large closed shape with two horizontal roll handles on belly and one vertical band handle to rim	horizontal roll handle, attached to body
cooking pot	cooking ware, medium closed shape with wide neck, vertical band or roll handle to rim	vertical handle, attached to rim or body
tripod cooking pot	cooking ware, neckless medium shape with rounded base and three oval or band-shaped feet attached on underside	tripod foot, attached to body (3 feet to one vessel)
pithos	heavy coarseware, very large closed shape, variety of rims and bases, number of handles unknown	

Fig. 12. Vessel types common at Kalapodi; diagnostic vessel parts.

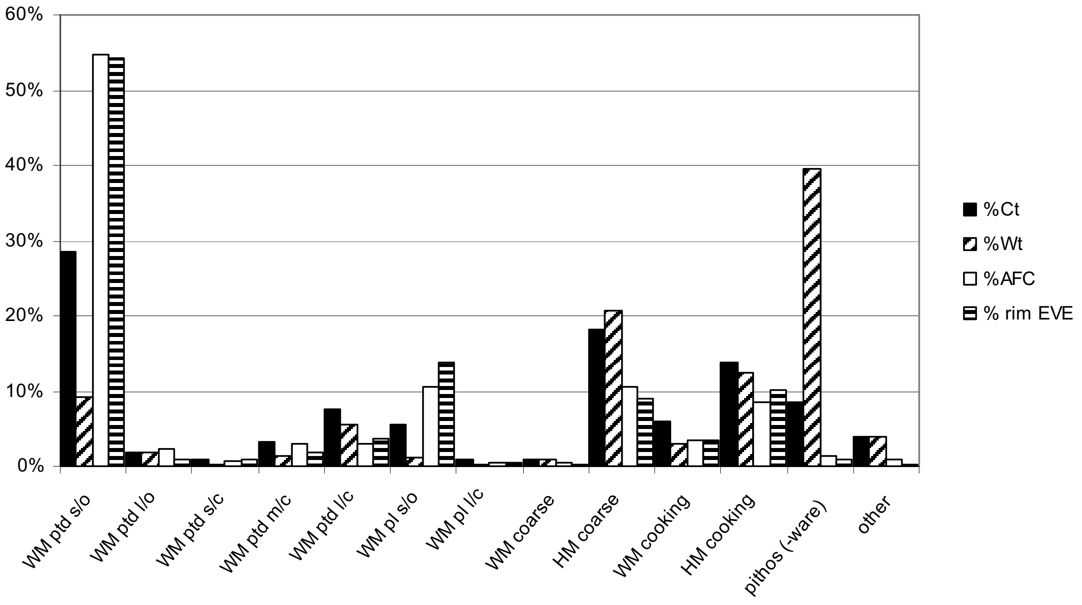

Fig. 13. Kalapodi trench 5030/4965. Overall sherd count – comparison of results from different quantitative methods.

All Lots	Feature sherds per vessel unit
WM ptd s/o	28.6
WM ptd l/o	50.6
WM ptd s/c	7.6
WM ptd m/c	34.8
WM ptd l/c	27.1
WM pl s/o	32.1
WM pl l/c	52.6
WM coarse	19.1
HM coarse	30.6
WM cooking	22.8
HM cooking	49.2
pithos (-ware)	24.2
average	31.6

Fig. 14. Average number of features per vessel (total feature count/rim+base EVE).

Phase	I		II		III		IV	
	%Ct	%Wt	%Ct	%Wt	%Ct	%Wt	%Ct	%Wt
WM coarse	18.6%	19.2%	0.2%	0.2%	0.0%	0.0%	0.7%	0.9%
HM coarse	3.3%	4.5%	24.8%	18.0%	24.1%	26.7%	33.0%	44.6%
HM cooking	4.9%	6.4%	22.3%	12.5%	29.4%	23.2%	3.6%	4.2%
pithos	5.4%	25.8%	26.6%	56.8%	6.8%	29.9%	2.1%	18.3%
HM overall	13.6%	36.7%	73.7%	87.3%	60.3%	79.8%	38.7%	67.1%

Fig. 15. Kalapodi phases I-IV, comparison of sherd count and weight for WM and HM coarsewares (pithoi are handmade throughout).

Phase	I		II		III		IV	
	%AFC	%rimEVE	%AFC	%rimEVE	%AFC	%rimEVE	%AFC	%rimEVE
WM coarse	11.3%	9.2%	0.8%	2.5%	0.0%	0.0%	0.2%	0.4%
HM coarse	1.2%	0.6%	18.0%	25.0%	13.7%	7.9%	15.4%	16.0%
HM cooking	4.0%	2.9%	21.9%	27.6%	17.0%	22.9%	1.9%	2.2%
pithos	0.3%	0.0%	7.0%	12.2%	1.4%	0.0%	0.4%	0.0%
HM overall	5.5%	3.5%	46.9%	64.8%	32.1%	30.8%	17.7%	18.2%

Fig. 16. Kalapodi phases I-IV, AFC and rim EVE for WM and HM coarsewares.

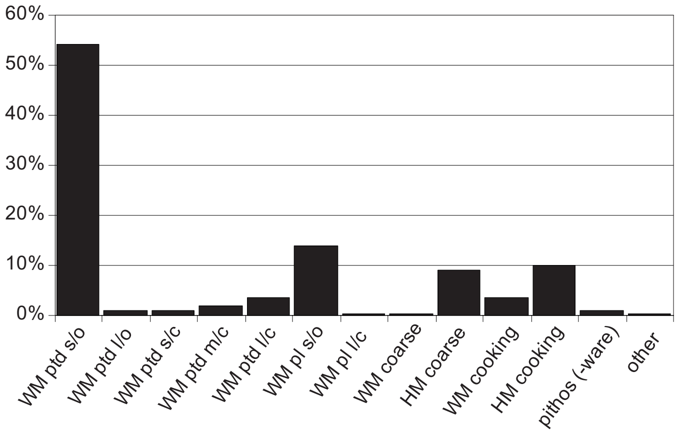

Fig. 17. Kalapodi trench 5030/4965. Summary of rim EVEs for trench. Representation of ceramic categories.

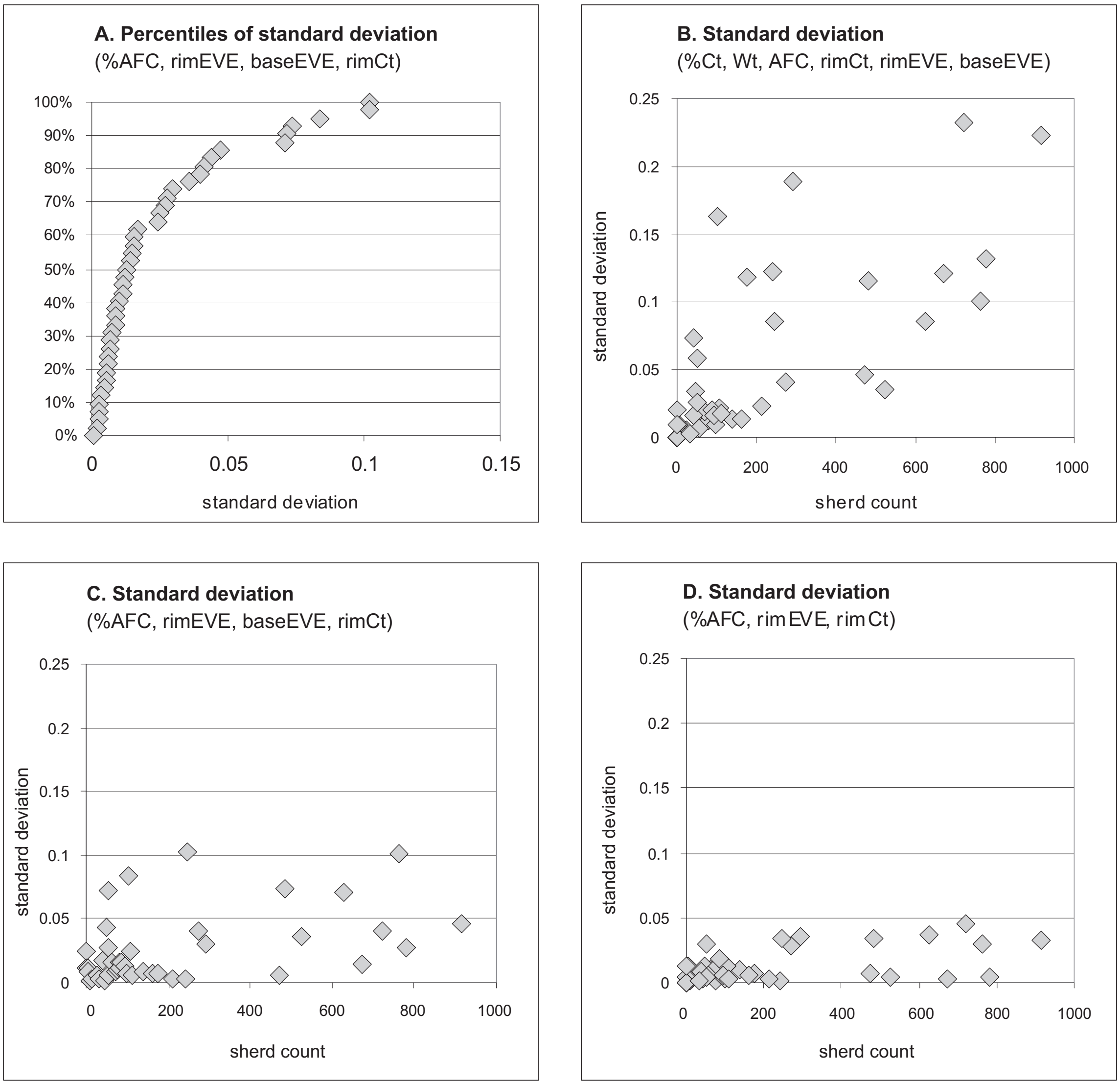

Fig. 18. Standard deviation, based on the percentages of different quantitative methods.
A: %AFC, rim EVE, base EVE, and rim count, ranked and shown in percentiles.
B-D: correlation between sherd count and standard deviation, for different sets of quantitative methods.

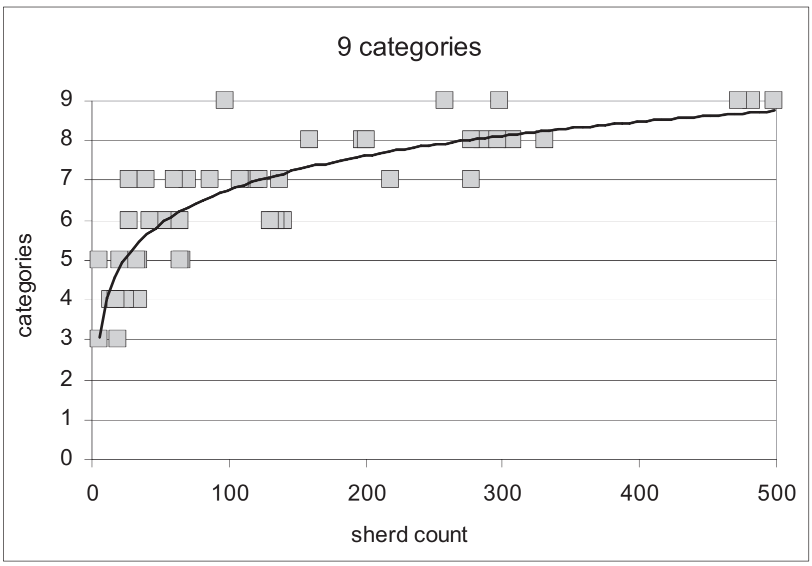

Fig. 19. Sample size (x-axis: sherd count per Befundnummer*) and representation of ceramic categories (y-axis).*

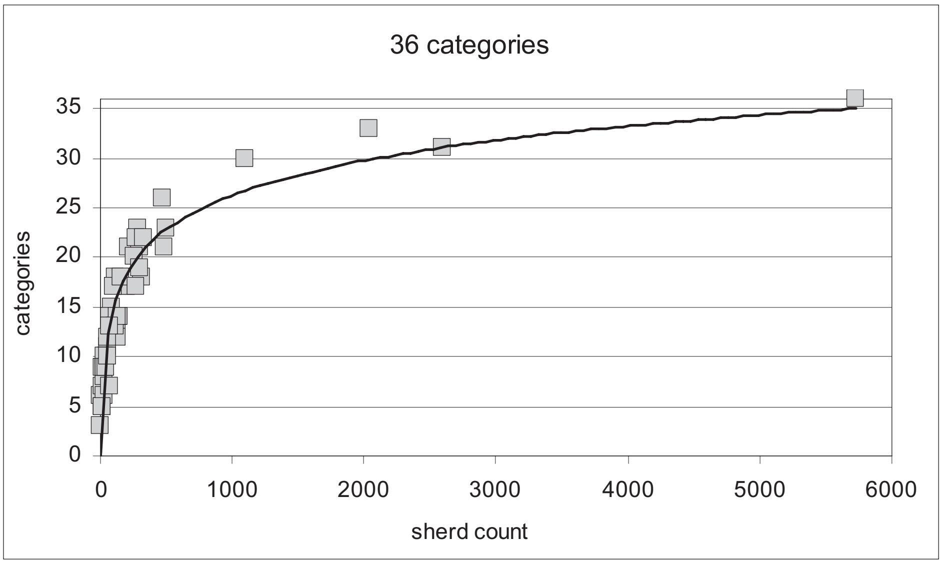

Fig. 20. Sample size (x-axis: sherd count per Befundnummer*; values above 1000 are for phases II, III, IV, and II-IV combined) and representation of ceramic categories (y-axis).*

	Phase II				Phase III				Phase IV			
	R (Ct)	H (Ct)	F (Ct)	B (Ct)	R (Ct)	H (Ct)	F (Ct)	B (Ct)	R (Ct)	H (Ct)	F (Ct)	B (Ct)
WM ptd s/o	15	14	12	61	104	54	49	515	116	115	134	551
WM ptd l/o	1	0	1	4	1	0	8	39	7	1	8	45
WM ptd s/c	0	0	1	0	1	0	0	4	2	1	1	21
WM ptd m/c	2	2	0	36	6	7	5	88	5	13	8	112
WM ptd l/c	1	3	3	79	1	0	4	74	3	3	4	69
HM coarse	11	13	8	240	11	22	16	578	26	24	29	592
HM cooking	11	16	1	217	31	21	9	702	3	7	0	64
pithos (-ware)	7	0	2	283	1	1	3	172	0	0	2	41
other	3	1	0	50	3	0	0	69	1	1	0	24
TOTAL	51	49	28	970	159	105	94	2241	163	165	186	1519

Fig. 21. Sherd counts for phases II-IV (Late Protogeometric to Middle/Late Geometric).