SOME PROBLEMS IN THE QUANTITATIVE ANALYSIS OF VEGETABLE REFUSE ILLUSTRATED BY A LATE HORIZON SITE ON THE PERUVIAN COAST

Mark Nathan Cohen

As archaeologists have shifted their attention from the identification of prehistoric events to an analysis of prehistoric processes, there has been a concomitant shift in refuse analysis away from the mere identification of animal and plant taxa toward an identification of the implications of preserved remains in terms of prehistoric diet and economic activity. Among the most significant such trends has been the recent stress on the quantitative evaluation of organic remains in the attempt to assess gradual shifts in diet or economic activity. The study which perhaps best illustrates this new approach to organic remains is the Tehuacan project in Central Mexico where quantitative shifts in diet and economic activity are defined with apparent great precision.[1]

The coast of Peru, where organic preservation is generally considered to be extremely good, offers an excellent laboratory for the study of prehistoric middens. Indeed, the consistency of organic preservation on the Peruvian coast suggests that it might well be one of our best laboratories for the study of dietary and economic shifts throughout the development of complex societies over a wide geographic area. It is my intention, however, to suggest that there are limits, as yet unrecognized, in the quantitative analysis even of "well preserved" middens such as those of the Peruvian coast, and that our analyses must, therefore, be much more cautious and limited in their conclusions than the Tehuacan work suggests.

Between 1968 and 1970, I undertook the analysis of organic remains from a series of archaeological sites from the central coast of Peru with the idea of emulating the work of MacNeish in quantitative analysis.[2] I was forced to conclude that, at least for the Peruvian coast, and probably for other regions as well, accurate quantitative measurement of organic remains is impossible; that only the most general of quantitative trends can be accurately envisioned; and that even those general trends can be reconstructed only tentatively until they have been reconfirmed by a large number of parallel lines of evidence.

This skepticism results in large part from the analysis of a single Late Horizon site in the Lurîn Valley (PV48-35). The midden from this site is such that it casts serious doubt on the quality of preservation at other sites, even where preservation is presumed to be excellent (for example, the famous Necropolis at Ancón, PV45-1, where I have worked).

PV48-35 is located on the southern margin of the Lurîn Valley, about 25 km. upstream from the coast. The site is a large one, (dimensions unspecified) with a number of preserved structures. The site was excavated in 1967 by Jane Feltham, then a graduate student at

Columbia University, as part of the Lurîn Valley survey undertaken by Thomas C. Patterson and students from Yale University and the University of California, Berkeley.[3] Feltham kindly made her organic refuse available for me to study. I undertook the sorting and identification of the vegetable material. Special problems in plant identification were referred to Margaret Towle or to specialists in the various plant taxa. The material came from a single cut 2 meters long by 1.5 meters wide and 1.5 meters deep (a total of about 4.5 cubic meters) representing a very small fraction of the total site. Feltham's analysis of the pottery indicated that the entire cut represented a Late Horizon occupation, hence representing a period of no more than 70 years. Feltham had divided her excavation into eight stratigraphic units. Because of the great quantity of refuse involved, I analyzed only between one-half and one-third of her material (a total of about 2 cubic meters), arbitrarily picking four of her eight excavation units for analysis.

A complete listing of the materials identified is shown in Table 1. For present purposes it is sufficient to note that this single cut, representing an accumulation of no more than 70 years, produced between 40,000 and 50,000 countable items of vegetable refuse alone. (The faunal remains had been bagged separately and were lost.) The total included such items as 6300 whole or partial maize cobs (Zea mays), 5500 valves of the pods of common beans (Phaseolus vulgaris), and nearly 25,000 cotton seeds (Gossypium barbadense) among a total of nearly 40 separate taxa of cultigens and wild plants. This total does not include such items as maize husk fragments and cotton fiber for which no meaningful counts could be obtained but which, in bulk, accounted for a sizeable fraction (30-40%) of the material studied. Since only half of the refuse was studied, the countable items of vegetable refuse from the excavated sample probably totaled on the order of 100,000 items. This is, I suggest, a reasonable approximation of "good organic preservation."

It should be noted that this figure is essentially comparable to the organic sample from the entire Tehuacan project. MacNeish cites a figure of 110,000 items of organic refuse recovered by that project.[4] The comparison between these two figures, in turn, raises serious questions about the definition of "significant samples" in archaeology. Reconstructing prehistoric diet or economy from preserved refuse is an exercise (explicit or implicit) in statistical inference; we infer total organic material consumed from a preserved sample. The accuracy of our inferences depends on the relative size of the preserved sample and the original quantity of organic material consumed (the universe from which the sample is taken). Until now, we have lacked a true appreciation of the volume of organic refuse built up over even a brief period of time, and therefore we have tended to underestimate grossly the size of the sampled universe. Reports which refer to samples of a few hundred or a few thousand items from a site, or tens of thousands of items from a sequence such as that of Tehuacan, have been presumed to represent significant samples from which reasonable inferences could be made. The excavation of the Lurîn Valley site, however, suggests that the total refuse items from a mere portion of a site of brief duration may well number in the tens of thousands if preservation is good (and even here, of course, we have no guarantee of total preservation). This in turn indicates that we must alter considerably our assumptions about what constitutes a "significant" sample. The reliability of our results should be reconsidered accordingly.

It is also instructive to compare the figures from PV48-35 with the counts of organic refuse which I obtained from a test excavation at the Ancón Necropolis (PV45-1), a site noted for its good organic preservation, where essentially the same range of organic remains was encountered. I undertook this excavation together with Eugene McDougle of Columbia University in 1969. Our excavations consisted of two cuts, each with a number of natural strata, all falling within the Late Intermediate Period. We excavated a total of about twelve cubic meters. The refuse contained only about 3000 countable vegetable items.

Why should there be such a difference in quantitative preservation between PV48-35 and PV45-1 when both sites display very much the same range of preserved organic materials? We are accustomed to assuming that the range of preserved plant remains is a good index of the quality of preservation. Comparisons between the two sites suggests that this assumption may be unjustified. Part of the difference in the quantity of preserved remains at the two sites may be due to the rate of accumulation of inorganic matrix. The more mineral sediments deposited, the less the apparent density of organic remains. It is also possible that, since the Necropolis is located directly on the shore of the ocean, vegetable foods may actually have been of secondary importance compared to seafoods, and this may account partially for the scarcity of plant material. The excavations at the Ancón Necropolis turned up shellfish at an average density of about 4000 per cubic meter. No comparable figures are available for the Lurín Valley site where we only know from Feltham's field notes that shell refuse was fairly abundant. I doubt, however, whether either of these factors can account entirely for the observed disparity in the density of preserved vegetable remains. We are faced with a dilemma. Apparently organic preservation can be excellent in the sense that a wide range of animal and plant materials is preserved, while at the same time, only a tiny fraction of the total vegetable consumption is represented. The explanation, I believe, is that plant material is highly sensitive to conditions of deposition such as soil chemistry or the immediacy with which the refuse, once deposited, is covered up. Given the right combination of these factors, one day's garbage may be preserved more or less intact and the next day's destroyed. The result would be an extremely fine alternation of preserved and unpreserved vegetable refuse which might display the full range of materials with no obvious signs of poor preservation and yet be quantitatively a poor sample. The only clue to this type of preservation would be the relatively low density of preserved vegetable remains per unit of volume. The conditions mentioned, I believe, constitute the reason for the relatively small number of plant remains obtained from the Necropolis.

One result of this sort of preservation is that the apparent ratios of plant and animal refuse in a midden may be misleading. Animal bone and shell, being relatively resistant to decay, may be preserved more or less continuously while vegetable remains are preserved only sporadically. If so, the result could be a situation where vegetable food remains are significantly under-represented relative to faunal remains even though the full range of vegetable materials is represented. This type of preservation characterizes a number of sites on the central coast in my experience (PV45-1, -25, -59, -100, -136, PV46-3), and I suspect that it is widespread. Eugene McDougle has noted similar scanty, though qualitatively wide-ranging, preservation at a number of other sites, also on the central coast, and has offered the further observation that there is a fairly consistent difference between the quantity of preservation in Late Horizon sites and that in sites of all preceding periods including the Late Intermediate Period.[5] Neither he nor I can explain this difference. We find no obvious differences in the environments of the different sites. Greater antiquity can hardly be a significant factor in distinguishing deposits of the Late Intermediate Period from those of the Late Horizon. I can only conclude that cultural differences in the treatment of refuse distinguish Late Horizon sites from those of other periods, at least on the central coast, and that these differences account for the variation in preservation. The important conclusion, however, is that even where organic preservation appears to be complete, as evidenced by the range of preserved materials, it may represent only an insignificantly small sample of the total original refuse. Statistical inferences should be tempered correspondingly.

There are a number of further problems in quantitative midden analysis which have tended to go unrecognized in recent work but which can also be illuminated by reference to the refuse from PV48-35. Paradoxically, these problems have developed because of the historical derivation of midden analysis techniques from studies directed primarily toward faunal analysis. Bone (or shell) has certain properties which are utilized in the existing analytic methodology and which have given rise to a number of analytic assumptions.[6] These assumptions are questionable even in the context of faunal analysis; when the same assumptions are applied to the full range of preserved refuse they lead potentially to serious misinterpretations.

The first significant property of faunal material is the regularity of its chemical structure. Whether bone or shell is being analyzed the chemical components of the hard parts of the various taxa are comparable. Hence the assumption is made that preservation is a constant. Bone (or shell) is either preserved in a certain locality of a site or it is not and, if preserved, it is presumed to be preserved unselectively; differential preservation is largely discounted.

The second property of bone (or shell) is that there is a high degree of regularity in the vertebrate (or invertebrate) body plan. There is a fixed and predictable quantitative relationship between the soft (and potentially edible) body parts and the hard, inedible (and potentially preservable) portions of the body. From this fact, the assumption has developed that there is a fixed and predictable relationship between preserved archaeological refuse and the calories consumed by the prehistoric population. If meat is eaten, bone or shell refuse is assumed to be left behind in predictable amounts.

Differences in butchering and preparation techniques are assumed to be recognizable and correctable since the regularity of the vertebrate skeleton allows us to recognize which pieces are not represented. (One has, however, no way of knowing how many animals are totally unrepresented in the refuse or how much of those animals represented was actually eaten.) The consequences of these two properties of faunal material and the assumptions that have been derived from them is that we have become accustomed to considering faunal refuse (and by extension, all refuse) as a fairly reliable indicator of the relative importance of various foods in the diet. One simply counts the number of preserved bones in each taxon, corrects for observed differences in preparation or butchering, counts the minimum or relative number of whole organisms of each taxon represented, corrects for their relative body weight and nutritive value, and establishes the relative number of grams of flesh or calories represented by each taxon. This methodology involves a number of untested, and questionable, assumptions which render even quantitative faunal analysis suspect. However, both the underlying assumptions and the methodology have recently been extended to vegetable remains with the result that quantitative estimates of total diet have appeared in the literature. In Tehuacan, for example, MacNeish offers estimates of the importance of various (animal and vegetable) items in the diet at different time periods based on the analysis of preserved middens in accordance with the type of methodology outlined above.[7]

I suggest that such estimates are invalid. The sources of error, which are relatively minor as long as one confines oneself to faunal material, become much greater when the analysis is extended to vegetable materials which lack both the chemical regularity and the anatomical regularity on which this methodology is based. Let us consider again the sample from PV48-35. Taken at face value, the refuse sample from this site suggests a number of fairly startling conclusions about the relative importance of various items in the diet. Not surprisingly, maize (Zea mays) and common beans (Phaseolus vulgaris) are among the most important food resources. Lima beans (Phaseolus lunatus) and squash (Cucurbita spp.) however are strikingly scarce in contrast to our general assumptions about the importance of these crops in the prehistoric Peruvian diet. It is also noteworthy that tubers of all types are very scarce. Highland tubers such as potatoes (Solanum tuberosum) and oca (Oxalis tuberosa) are almost unknown at this coastal site, a fact which is perhaps not too surprising. What is surprising is that such good low altitude root crop staples as sweet potatoes (Ipomoea batatas) and manioc (Manihot esculenta) are only sparsely represented. Even more surprising is the fact that the only tuber represented in abundance is achira (Canna sp.) usually considered to be a minor root crop. It is also interesting that lucumas (Lucuma bifera), a minor fruit used today mostly for ice cream, are fairly abundant, while presumably more common fruits such as guavas (Psidium guajava) and avocados (Persea americana) are scarce. These facts may call for some redefinition of our assumptions about Late Horizon economic patterns on the central coast if they are taken as representative of true consumption. On the other hand, quantitative trends in the sample quickly lose their significance if we analyze some of the factors that govern the production and preservation of vegetable refuse.

In the first place, the preservation patterns of vegetable refuse are much more complicated than those of bone or shell. Unlike meat and bone or shell, the edible portions of plants and the associated inedible portions, which remain as refuse, represent a wide variety of different biological structures with different durability. Whereas, bone or shell are presumed to be preserved on a more or less all-or-none basis, the presence of some plant material, or even a significant variety of plant remains, does not guarantee that all types of plant refuse are preserved equally. Clearly, maize cobs, silica-bearing grasses, and seed structures designed to withstand weathering and rotting will be more easily preserved than soft items such as starchy tubers. Hence, there is no way of assuring that the relative quantities of preserved plant remains represent the original refuse content, even if a wide range of vegetable materials is preserved. This statement may seem self evident, but unfortunately, I believe, the tendency to regard vegetal preservation as analogous to the preservation of bone is fairly widespread. The problem is greatly complicated by the fact that there appears to be no simple, clear-cut linear order of durability or "preservability" of plant structures. Except in a very crude sense, as stated above, we cannot predict from the general level of preservation at a site exactly which plant structures should have been preserved (whose absence is therefore significant) and which are likely simply to have been lost by decay. The order in which various taxa decay appears to depend on soil types and moisture conditions. The relative order of decay might be predictable for certain conditions provided a large enough sample of sites could be studied to eliminate the interference of other variables, but such studies have not been done. This problem may be confusing the interpretation of refuse even at PV48-35 where organic preservation is apparently so complete. We are not as yet in a position to determine whether the scarcity of tubers reflects their poor preservation or actual lack of use. Nor do we know whether the relative abundance of lucumas reflects their use or the durability of the seeds.

Even if we could discount differential preservation, the problem of dealing with vegetables is complicated by the problem presented by different patterns of utilization. The various vegetable foods with their various structures can be prepared in entirely different ways affecting the contents of the resulting refuse. Since vegetable foods are more readily stored and transported before eating than are animal foods, differences in preparation can have a marked effect on the content of a given site. For example, Lawrence Kaplan has pointed out that when beans are eaten their representation in the refuse depends entirely on whether or not they are shelled (and the pods discarded) at the site or elsewhere.[8] Cutler and Whitaker have suggested that squash remains may not appear in refuse if they were eaten green (and eaten whole) or were cut into strips and dried in the field before being brought back to camp.[9] The scarcity of both lima beans and squash at PV48-35 may be a simple function of such patterns of utilization. The absence of tubers may simply indicate that they were eaten with their skins.

A further problem stems from the methods employed in plant taxonomy and may not be appreciated by archaeologists who have never worked with botanical classification. The quantitative sample may reflect errors of classification favoring certain taxa. In faunal analysis this problem is less significant, because animal skeletons or shells are more or less comparable in the degree to which various parts are diagnostic of the particular taxon. Perkins has even been able to work out a correction factor to account for slight differences in the numbers of identifiable elements in the skeletons of different animals.[10] The range of structures involved in plant materials, however, makes this kind of comparison difficult. Classification of plants is based on flowers, which are rarely well preserved and rarely constitute the important economic portion of the plant. Success in identification depends on the degree to which nonflowering parts can be identified. I suggest that this will skew samples in two senses. First, domestic plants of economic importance today are much more thoroughly studied than are wild plants or plants of little contemporary economic significance. Wild plants and ones of little contemporary importance, therefore, will almost always be underrepresented in the identified refuse. In addition, the identified refuse may be skewed in favor of those taxa with highly distinctive characteristics easily recognizable when encountered in small fragments. In the sample from PV48-35, the relative abundance of lucuma may be due not only to the durability of the seeds but also to the fact that the seeds coats are a distinctive shiny brown which is readily recognizable even in small fragments. Similarly, achira may be the only tuber recognized in abundance because of its distinctive annular skin texture.

There are, moreover, further sampling errors inherent in the analysis of vegetable remains. The recovery of bone is amenable to fast excavating techniques and large scale excavations. Accurate recovery of plant remains, which are usually smaller and more delicate than bone, requires more careful work and in fact is extremely tedious when preservation is good. Without great expenditure of time and money, quantitatively complete recovery of plant remains is inconsistent with large scale excavation. For the study of plant remains, small test pits such as that employed by Feltham must be employed, and the refuse must be carefully screened or flotation must be used. A number of carefully spaced test pits may help to randomize the sample, but since both open sites and cave sites are often spotty in their preservation, it is not always possible to obtain a random effect. Many of the sites I excavated on the central coast had good preservation over a tiny fraction of their extent. It is becoming a truism in archaeology that artifact distribution reflects specialized zones of activity within the bounds of a single site, and the same is presumably true of organic refuse; thus, the sample recovered from an individual test pit may not be a representative indication of economic activity at the site. The extremely high ratio of common beans to lima beans and the absence of squash at PV48-35 may reflect nothing more than the economic specialization or dietary preferences of a single family. The degree of sampling error becomes even greater if, as at Tehuacan, samples are taken to be representative of an entire region rather than just a portion of a particular site.

Thus, such "patterns" as the absence or scarcity of squash, lima beans, and most tubers, or the abundance of lucuma and achira at PV48-35 are hardly to be considered indicative of economic patterns for the Late Horizon in the Lurîn Valley or even for that site as a whole. What are represented are isolated examples which should be noted for comparison with other excavations. Their actual significance is a function of the regularity with which similar patterns can be found to repeat themselves in comparable sites.

In sum, I do not believe that reliable quantitative estimates of diet can be made from direct counts of items of food refuse. It is not my intention to suggest that we eliminate quantitative analysis which, in many ways, is essential to sophisticated modern studies. What is suggested by my material is that caution must be employed in the use of quantitative data. We must soften the focus of our quantitative studies and recognize that while data may reflect gross economic trends they probably cannot measure these trends very accurately. Furthermore, quantitative conclusions must be considered highly tentative except where they reflect gross trends observed over many large samples; and we must consider alternative explanations, including differential preservation and variations in preparation in evaluating our data even where preservation is "excellent."

NOTES

[1]MacNeish, 1967.

[2]Cohen, ms.

[3]Feltham, field notes.

[4]MacNeish, 1967, p. 290.

[5]McDougle, personal communication.

[6]Daly, 1969; Perkins, 1964; Perkins and Daly, 1968.

[7]MacNeish, 1967.

[8]Kaplan, personal communication.

[9]Cutler and Whitaker, 1961.

[10]Perkins, 1969.

BIBLIOGRAPHY

Cohen, Mark Nathan
ms. Population growth, subsistence, and settlement in the Ancón-Chillón region of the central coast of Peru. Ph.D. dissertation, Department of Anthropology, Columbia University. New York. 1971.

Cutler, Hugh Carson, and Whitaker, Thomas W.
1961 History and distribution of the cultivated Cucurbits in the Americas. American Antiquity, vol. 26, no. 2, April, pp. 459-485. Salt Lake City.

Daly, Patricia
1969 Approaches to faunal analysis in archaeology. American Antiquity, vol. 34, no. 2, April, pp. 146-153. Washington.

MacNeish, Richard Stockton
1967 A summary of subsistence. The Prehistory of the Tehuacan Valley, vol. 1, edited by Douglas Byers, pp. 290-309. University of Texas Press, Austin.

Perkins, Dexter Jr.
1964 The prehistoric fauna from Shanidar, Iraq. Science, vol. 144, no. 3626, June, pp. 1565-1566. Washington.

1969 Fauna of Çatal Hüyük: Evidence for early cattle domestication in Anatolia. Science, vol. 164, no. 3876, April, pp. 177-179. Washington.

Perkins, Dexter Jr., and Daly, Patricia
1968 The potential of faunal analysis: An investigation of the faunal remains from Suberde, Turkey. Scientific American, vol. 219, no. 5, May, pp. 96-106. New York.

TABLE 1

Identified Vegetable Remains from PV48-35

	Levels 3	5B	7	8
Zea mays				
cobs	500	1200	500	250
cob fragments	400	1000	1200	1275
kernels	-	1280	400	6
tassel fragments	P	P	P	P
husk fragments	A	A	A	A
Cucurbita sp.				
seeds	1	85	1	-
rind fragments	-	-	-	2
peduncles	3	12	16	3
Cucurbita moschata				
seeds	-	11	2	1
Cucurbita maxima				
seeds	-	3	-	-
Lagenaria sp.				
seeds	5	275	3	1
rind fragments	71	115	126	190
peduncles	-	3	-	-
Phaseolus vulgaris				
seeds	153	256	904	186
pod halves	320	2720	2400	165
Phaseolus lunatus				
seeds	3	7	4	-
pod halves	-	10	-	2
Canavalia sp.				
seeds	21	87	88	18
pod halves	2	1	-	2
Inga feuillei				
seeds	-	74	68	-
pod fragments	17	22	76	10
Galactia striata				
pod halves	-	1	-	-
Arachis hypogaea				
seeds	-	-	-	-
pod fragments	30	860	480	280

Levels	3	5B	7	8
Caesalpinia sp.				
seeds	-	12	-	13
pod fragments	-	-	-	-
Erythrina sp.				
seeds	-	3	-	10
pod fragments	2	1	-	1
Prosopis sp.				
seeds	-	-	-	-
pod fragments	-	-	-	-
Other legume				
seeds	4	20	-	-
pod fragments	-	5	2	-
Ipomoea batatas				
tuber fragments	2?	15?	2?	-
Manihot esculenta				
tuber fragments	-	8	2	-
Canna sp.				
tuber fragments	4	187	97	48
leaf fragments	-	5	27	16
Oxalis tuberosa				
tuber fragments	-	1?	-	-
Solanum spp.				
tuber fragments	-	1?	-	-
Other (wild) tuber fragments	-	9	6?	-
Bunchosia armeniaca				
seeds	17	95	75	7
Campomanesia lineatifolia				
seeds	-	2	-	-
Lucuma bifera				
seeds	64	374	163	110
Psidium guajava*				
fruits	1	26	13	2
Sapindus saponaria				
seeds	-	-	19	-

	Levels 3	5B	7	8
Persea americana				
seeds	3	12	7	1
Capsicum sp.				
fruits	-	6	-	-
stems	-	8	-	-
seeds	P	11	P	P
Capsicum baccatum				
fruits	-	11	-	-
Asclepias sp.				
pod fragments	-	1	5	1
Erythroxylon sp.				
seeds	6	24	40	-
Gossypium barbadense				
fiber	A	P	A	A
seeds	1700	A	18,400	4500
boll fragments	17	67	100	44
Other fiber	P	P	P	P
Schinus molle				
seeds	-	17	-	-
Unknown seed/fruit	-	6	-	-
Gramineae				
culm fragments	600	664	1600	1100

P = present
A = abundant
* No separate count was kept of P. guajava seeds.