



A Mathematical Model of the Effects of Screen Size on Zooarchaeological Relative Abundance Measures

Michael D. Cannon

Department of Anthropology, Box 353100, University of Washington, Seattle 98195, U.S.A

(Received 26 February 1998, revised manuscript accepted 19 May 1998)

Zooarchaeologists have long recognized that the use of different sizes of screen in excavation can result in varying degrees of bias in faunal samples. This paper presents a graphical, mathematical representation of the ways in which screen size choice can affect measures of taxonomic relative abundance. Data from the site of Mogollon Village in southwestern New Mexico show that the rank order of samples along a measure of relative abundance like the Artiodactyl Index can vary with screen size. This occurs because specimens of a given taxon can respond to changes in screen size in different ways in different provenience units, a fact which also severely limits the utility of “correction factor” methods for dealing with screen size bias. The model presented here and the Mogollon Village fauna have important implications for studies based on temporal or spatial trends in the abundances of large-bodied taxa relative to small-bodied taxa, especially in cases where large mesh sizes or no screens at all were used in excavation.

© 1999 Academic Press

Keywords: ZOOARCHAEOLOGY, ARCHAEOLOGICAL METHODS, SCREEN SIZE.

Introduction

As zooarchaeology has become increasingly quantitative over the last few decades, faunal analysts have come to recognize that techniques used in data collection can have substantial impacts on the results of zooarchaeological studies. In particular, many researchers have documented that the size of the mesh through which excavated sediments are screened can have dramatic effects on the kinds of taxa, numbers of taxa and relative abundances of different-sized taxa that are recovered (e.g., Butler, 1993; Casteel, 1972, 1976; James, 1997; Nagaoka, 1994; Payne, 1972; Shaffer, 1992; Shaffer & Sanchez, 1994; Thomas, 1969; see also Stahl, 1996: 33–34).

These previous works on the effects of screen size on faunal recovery have used experimental or archaeological data to evaluate empirically the biases that screen size choice can introduce. Here, I explore on a more conceptual level what effects the differential recovery of taxa will have on selected measures of relative abundance. The graphical, mathematical model presented here allows one to see how both the ordinal ranks of the abundances of different taxa within an assemblage, and the ordinal ranks of different assemblages along an index of relative abundance like the Artiodactyl Index, can vary as the result of screen size choice.

Faunal data from the Mogollon Village site in southwestern New Mexico will be used in conjunction with this model to illustrate how differential recovery can cause errors in ordinal scale measures of relative

abundance. This assemblage also demonstrates that the assumptions required by the “correction factor” method for dealing with screen size bias (Thomas, 1969; James, 1997) might not be met in many cases, and that the use of “correction factors” can result in abundance values which bear little resemblance to values actually observed.

Relative abundance measures and screen size

Measures of the relative abundances of different-sized taxa in faunal assemblages can be useful for a variety of purposes. They have become particularly important in recent attempts by archaeologists working in western North America to document cases of prehistoric resource depression (Charnov, Orians & Hyatt, 1976), or reductions in the prey capture rates of prehistoric peoples as a direct result of their own activities. By examining temporal declines in the abundances of high energetic return large-bodied taxa relative to lower return smaller taxa, a number of authors have argued that resource depression occurred in such areas as California (Broughton, 1994a, b, 1995, 1997; Hildebrandt & Jones, 1992), the Great Basin (Grayson, 1991) and the southwestern United States (Szuter & Bayham, 1989; Harvey & Broughton, 1996; Janetski, 1997; see also Bayham, 1979; for a more detailed discussion of archaeological resource depression studies see Grayson & Cannon, 1999). Most of these studies have employed measures analogous to the Artiodactyl Index developed by Bayham (1982; Szuter

& Bayham, 1989), calculated as the ratio of artiodactyls in an assemblage relative to the total number of artiodactyls and leporids in that assemblage ($\Sigma \text{Artiodactyls}_i / \Sigma [\text{Artiodactyls}_i + \text{Leporids}_i]$). This is only one of many possible such pairs of large and small taxa, however, and other indices have also been used (e.g., Broughton, 1994a, b, 1995, 1997).

As Grayson (1984) has argued, if these kinds of relative abundance measures are based on NISP values, they are subject to the problem of interdependence among the things that are being counted and, as a result, should not be treated as ratio or interval scale measures. If they are based on MNI values the interdependence problem disappears, but one is then faced with the even more troublesome fact that the measures can vary depending on how the samples are aggregated. The question thus becomes: are these kinds of measures, when based on NISPs, valid even on an ordinal scale?

Previous works on the effects of screen size have shown that it is generally the case that specimens of larger-bodied taxa are recovered in higher proportions (relative to the number of specimens present in the ground at the time of excavation) than are specimens of smaller-bodied taxa. It is thus indeed reasonable to ask whether measures of relative abundance that incorporate different-sized taxa really are valid on an ordinal level when screens are used in excavation (as opposed to when recovery is complete). Below, I demonstrate how differential recovery levels among taxa can interact with screen size choice in such a way that the ordinal ranking of the relative abundances of taxa within a sample, or of different samples along some index of relative abundance like the Artiodactyl Index, can vary depending on the size of mesh used. The possibility that these kinds of effects are present must be addressed in any analysis based on relative abundance measures, especially when the taxa involved are of such different sizes that the proportions of them recovered in a given size of screen are likely to be very different.

The Differential Recovery Model

Changing rank order of relative abundances within an assemblage

Here I present a model of differential taxonomic recovery that can be used to explore how screen size choice can introduce systematic errors into both the rank order of the relative abundances of different taxa within a sample and into the rank order of multiple samples along some index of relative abundance. The model deals only with cases in which two taxa are involved partly for the sake of simplicity, but also because faunal resource depression studies have been based on ratios incorporating two taxa. When more than two taxa are involved matters become more complex, though the implications of the model for such

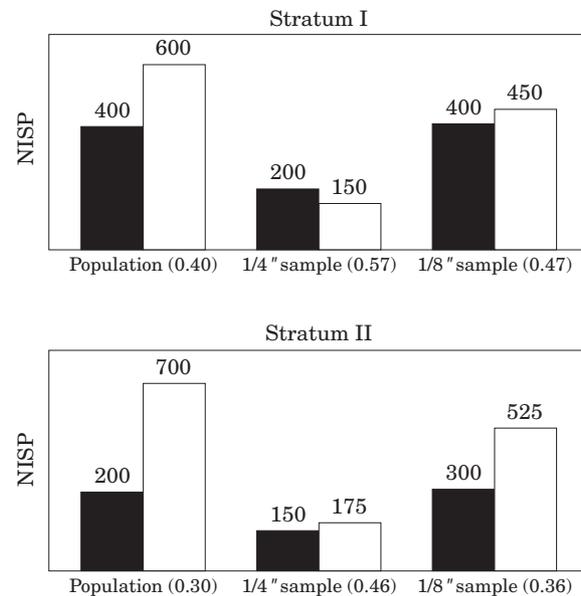


Figure 1. Population and sample NISP values for two taxa from two strata in an imaginary site in which 1/4" screen recovers 50% of Taxon A and 25% of Taxon B, and 1/8" screen recovers 100% of Taxon A and 75% of Taxon B. Shown in parentheses are the values of each population and sample along an index calculated as $\Sigma A_i / \Sigma (A_i + B_i)$. ■, Taxon A; □, Taxon B.

cases are not developed here. I start with the simpler problem of the rank order of the relative abundances of taxa within a sample, and before introducing the model itself consider a hypothetical example.

An imaginary site (Figure 1) contains two strata, I and II, with I overlying and thus younger than II. Each stratum contains two taxa, A and B, in different proportions. If all the identifiable bones in each stratum could be recovered, an index calculated as $\Sigma A_i / \Sigma (A_i + B_i)$ would give values of 0.40 for Stratum I and 0.30 for Stratum II. For the purposes of this example, Taxon A is the larger-bodied of the two taxa and will be more completely recovered no matter what size of mesh is used in excavation. I initially assume that the specimens of each taxon have been fragmented to the same degree in both strata, resulting in the recovery of equal proportions of identifiable specimens for each taxon between the two strata. Screening excavated sediments through 1/4" mesh will recover 50% of Taxon A and 25% of Taxon B, while 1/8" mesh will recover 100% of Taxon A and 75% of Taxon B.

It can be seen immediately that if only 1/4" screen were used at this site, the rank order of the relative abundances of the two taxa in Stratum I would be reversed with respect to the rank order of their relative abundances in the population being sampled. If 1/8" screen were used, however, their rank order in the resulting Stratum I sample would correctly reflect their rank order in the population. There is nothing particularly novel about this insight; many earlier studies on screen size biases have demonstrated empirically that

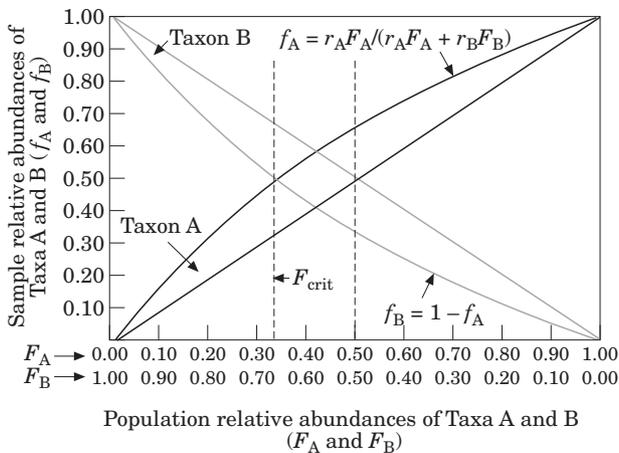


Figure 2. The differential recovery model.

this kind of thing can occur (e.g., Nagaoka, 1994). What is interesting to note, though, is that while the 1/4" sample from Stratum I provides incorrect results, the 1/4" sample from Stratum II does not.

Figure 2, the differential recovery model, shows how such a switch in the rank order of taxonomic abundances can occur as the result of incomplete recovery of specimens. The horizontal axis in this figure represents the range of all possible values that the relative abundance of either taxon can take in a population consisting only of Taxa A and B; the relative abundance of Taxon B, of course, is just 1 minus the relative abundance of Taxon A. The vertical axis represents the relative abundance of each taxon that would be seen in a sample collected in a given size of screen. The relative abundance of either of the two taxa in a screened sample is a function of three things: the relative abundance of that taxon in the underlying population being sampled, the proportion of it recovered in the screen size being used and the proportion of the other taxon recovered in that screen size. Let f_A and f_B represent the relative abundances of each taxon observed in a screened sample, let F_A and F_B represent the relative abundances of the two taxa in the population, and let r_A and r_B represent the proportion of each taxon that is recovered in a given size of screen. The sample relative abundance of each taxon can then be calculated as:

$$f_A = r_A F_A / (r_A F_A + r_B F_B), \quad (1)$$

and

$$f_B = r_B F_B / (r_A F_A + r_B F_B) = 1 - f_A. \quad (2)$$

The solid straight lines in Figure 2 represent $f_A = F_A$ and $f_B = F_B$, or the situation in which the relative abundances of the two taxa in a sample would perfectly correspond to their values in the population being sampled. This will occur either when a sample achieves complete recovery of specimens or when the taxa are recovered in equivalent proportions even though both

are only partially recovered. This can be seen by solving for equation (1) when $r_A = r_B$:

$$\begin{aligned} f_A &= r_A F_A / (r_A F_A + r_B F_B) \\ &= r_A F_A / (r_A (F_A + F_B)) \\ &= F_A / (F_A + 1 - F_A) = F_A. \end{aligned} \quad (3)$$

The curved lines in Figure 2, which we can call *sampling functions*, represent what will happen when the two taxa are recovered in different proportions. If Taxon A is recovered in a higher proportion than Taxon B in a given sized screen, that is if $r_A > r_B$, then Taxon A will be overrepresented in a screened sample with respect to its true relative abundance in the population and Taxon B will be underrepresented (such that $f_B = 1 - f_A$), except for the trivial case of when the population relative abundance of either taxon is zero. I am assuming here that r_A is greater than or equal to r_B ; obviously, if r_A is less than r_B then the curves in Figure 2 simply become a mirror image of the way in which they are portrayed here and the logic of the model still applies.

When will an error in the rank order of the relative abundances of the two taxa within a sample occur? Examine the two curved sampling functions in Figure 2 and note that the point at which they intersect is different from the point at which they would intersect if the samples were perfectly representative of population relative abundances. In other words, the point along the population relative abundance scale at which f_A becomes greater than f_B is skewed with respect to the point along this scale at which F_A becomes greater than F_B (which occurs, of course, when the population relative abundance of both taxa equals 0.50). The value of F_A at which f_A becomes greater than f_B can be termed F_{crit} . Samples taken from populations that lie between $F_A = F_{crit}$ and $F_A = 0.50$ will give an incorrect representation of the rank order of the relative abundances of the two taxa in the population.

The value of F_{crit} can be solved for by setting $f_A = f_B$ (from equations 1 and 2): F_{crit} is where

$$r_A F_A / (r_A F_A + r_B F_B) = r_B F_B / (r_A F_A + r_B F_B),$$

or where

$$r_A F_A = r_B F_B,$$

or where

$$F_A = r_B F_B / r_A. \quad (4)$$

The exact values of F_A and F_B can never be known in a real-world case, of course, unless all of the specimens initially present in the ground are recovered. The point of this discussion, though, is to assess how accurate any estimates of these numbers might be, at least on an ordinal level, when it is known that recovery of the specimens present in the ground at the time of excavation is not complete (as is the case when screens are used). Neither can the exact values of r_A or r_B be known in a particular case unless one first screens

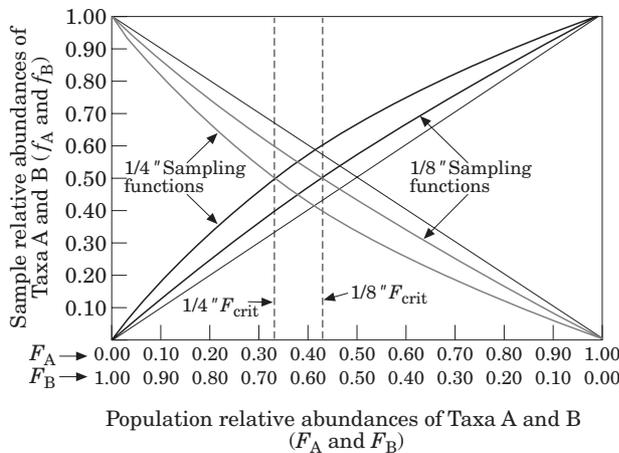


Figure 3. 1/4" and 1/8" screen sampling functions for the example presented in Figure 1.

sediments and then goes back and collects all of the missed specimens from the backdirt. Examination of equation (4), however, can allow us to predict mathematically how F_{crit} will respond to changes in the recovered proportions of the taxa involved, such as might occur due to changes in the size of mesh used in excavation.

F_{crit} can vary between $F_A=0.50$, which occurs when sample relative abundances are perfectly representative of population relative abundances, and $F_A \approx 0$, which occurs when Taxon B is missed in a sample altogether. (F_{crit} can never actually reach $F_A=0$ because in such a population there would be no specimens of Taxon A to be sampled.) Furthermore, F_{crit} is directly proportional to r_B and inversely proportional to r_A . As the ratio r_B/r_A approaches zero, F_{crit} approaches $F_A=0$, and when r_B/r_A approaches 1, F_{crit} approaches $F_A=0.50$. We can thus predict that F_{crit} will move closer to $F_A=0.50$ whenever r_B increases proportionately greater than does r_A with any change in recovery proportions.

This is illustrated by Figure 3, which presents the sampling functions in operation for the example in Figure 1. The change in screen size in this example resulted in a greater percentage increase in the recovered proportion of Taxon B than in the recovered proportion of Taxon A: r_A increased by a factor of 2 from 50% to 100%, while r_B increased by a factor of 3 from 25% to 75%, and the sampling functions became less biased accordingly. Figure 3 shows why the rank order of the relative abundances of the two taxa in Stratum I in this example is incorrect in the 1/4" sample but not in the 1/8" sample. For the 1/4" screen sampling functions the population relative abundance of Taxon A (0.40) falls between F_{crit} (0.33) and $F_A=0.50$, but for the 1/8" screen sampling functions it does not fall between F_{crit} (0.43) and $F_A=0.50$. We can also see why the samples from Stratum II show no such switch: the population relative abundance of Taxon A for this stratum (0.30) is such that it does not fall between

F_{crit} and $F_A=0.50$ for either screen size sampling function.

This imaginary example, of course, does not indicate that the use of 1/8" screen will always give the correct rank order of the relative abundances of taxa within a sample. Consideration of the model, however, allows us to see that the use of finer screen will result in a smaller range of population relative abundances within which errors will occur if the percentage increase in the recovered proportion of the less completely recovered taxon is greater than the percentage increase in the recovered proportion of the more completely recovered taxon.

Finally, and most important, note that in the example in Figure 1, even though the rank order of the relative abundances of the two taxa is incorrect in the Stratum I 1/4" screen sample, the rank order of the $\Sigma A_i/\Sigma(A_i+B_i)$ index between the two strata stays the same no matter what size of screen is used. The values of this index for each stratum vary greatly on an interval scale, to be sure, and are much closer to their true values in the 1/8" samples than in the 1/4" samples, but the index value of Stratum I always stays greater than that of Stratum II. I next consider how errors in the ordinal ranks of different samples along such indices of relative abundance can occur.

Changing rank order of assemblages along an index of relative abundance

I return to the example presented in Figure 1, but introduce the additional factor of differential fragmentation of specimens between the two strata. Figure 4 presents this kind of case: Taxon A is now more fragmented in Stratum I than it is in Stratum II, such that 1/4" screen will recover 30% of its identifiable specimens in Stratum I and 1/8" screen will recover 90% of them. To keep things simple I will keep Taxon A at its previous recovered proportions of 50% and 100% for the two screen sizes in Stratum II, and Taxon B at its previous recovered proportions of 25% and 75% for the two screen sizes in both strata. The rank order of the two strata along the $\Sigma A_i/\Sigma(A_i+B_i)$ index in the 1/4" screen samples is now incorrect. If only 1/4" screen were used at this imaginary site, one would see a temporal trend in the relative abundances of these two taxa which is, in fact, the opposite of what truly occurred.

Figure 5 shows how such a switch in the rank order of samples along an index of relative abundance can occur. This figure is analogous to Figure 2, but Taxon B has been left off for clarity. The 1/4" screen samples from the two strata in the example in Figure 4 are represented here. Introducing the factor of differential recovery of Taxon A between the two strata has resulted in them becoming subject to different sampling functions, which has, in turn, caused Stratum II to show a higher relative abundance of Taxon A than Stratum I.

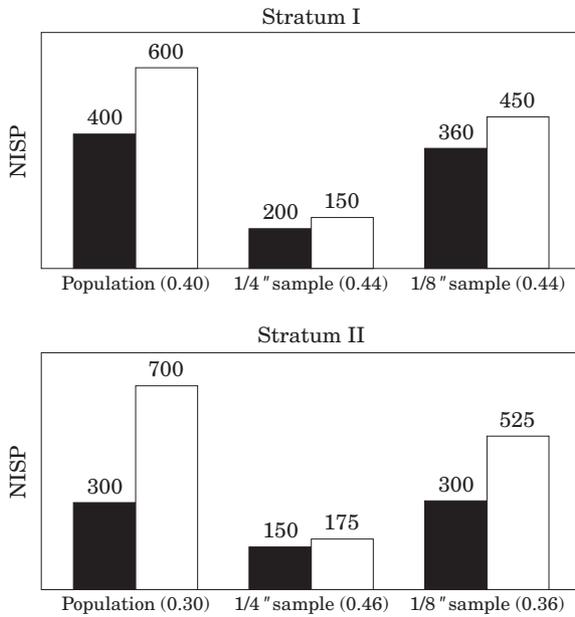


Figure 4. Population and sample NISP values for the two strata from Figure 1 when the recovered proportion of Taxon A is allowed to vary between strata. For Taxon A in this example, 1/4" screen recovers 30% in Stratum I and 50% in Stratum II, while 1/8" screen recovers 90% in Stratum I and 100% in Stratum II. For Taxon B, 1/4" screen recovers 25% and 1/8" screen recovers 75% in both strata. Quantities in parentheses are the values of each population and sample along the $\Sigma A_i/\Sigma(A_i+B_i)$ index. □, Taxon A; ■, Taxon B.

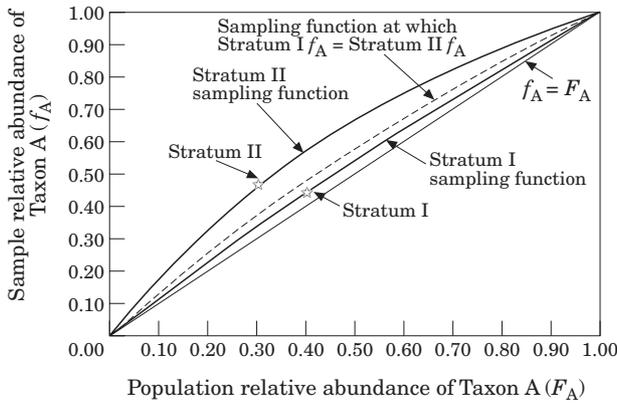


Figure 5. Sampling functions for the 1/4" screen samples in the Figure 4 example.

Recall from the previous section that the sampling functions are determined by the ratio r_B/r_A . As r_B/r_A approaches 1, a sampling function will approach linearity, producing samples that are perfectly representative of population relative abundances. As r_B/r_A decreases, a sampling function will become more biased. The dashed line in Figure 5 indicates the critical sampling function for this example (i.e., the critical value of r_B/r_A) at which the sample relative abundance of Stratum I would equal the sample relative

abundance of Stratum II. Whenever the r_B/r_A ratio of Stratum I is greater than this critical value, an error in the rank order of the two strata along an index of relative abundance will occur.

Solving for this critical sampling function, or the critical value of r_B/r_A , can, of course, never be done in the real world, nor could it ever be known whether the r_B/r_A ratio of the sample with the higher population relative abundance of Taxon A were greater than the critical r_B/r_A value. This would again require knowing the underlying population relative abundances and the values of r_A and r_B for both samples. Nevertheless, it is possible to use the graphical model in Figure 5 to explore the conditions under which errors in the rank order of two samples along an index of relative abundance will be more or less likely to occur.

First, such errors will only occur when the r_B/r_A ratio of the sample with the lower population relative abundance of Taxon A is lower than the r_B/r_A ratio of the sample with the higher population relative abundance of Taxon A. In other words, they can only occur when the sample from the population with the lower relative abundance of Taxon A also recovers a lower proportion of the specimens of Taxon B and/or a higher proportion of the specimens of Taxon A than does the other sample. This can be seen by considering the example in Figures 4 and 5. It is Stratum I in this example, not Stratum II, in which the recovered proportions of Taxon B are higher relative to those of Taxon A. (The value of r_B/r_A is 25/30 for Stratum I 1/4", 25/50 for Stratum II 1/4", 75/90 for Stratum I 1/8", and 75/100 for Stratum II 1/8".) If it were Stratum II which had a higher r_B/r_A ratio for the 1/4" samples, then, in terms of Figure 5, the sampling function for Stratum II would fall below and to the right of the Stratum I sampling function and there would be no way in which an error in the rank order of the two samples along the $\Sigma A_i/\Sigma(A_i+B_i)$ index could occur.

What if the recovered proportion of Taxon B were also allowed to vary between the two strata in this example? Again, errors in the rank order of the samples along an index of relative abundance will only occur when the sample from the population with the lower relative abundance of Taxon A has a lower r_B/r_A ratio than does the sample with the higher population relative abundance of Taxon A.

How much lower the r_B/r_A ratio of the sample from the population with the lower relative abundance of Taxon A must be in order for errors to occur is determined by the degree of difference between the population relative abundances. A corollary to this is that errors are more likely to occur the more similar the underlying populations are in terms of relative abundance. If, for instance, the Stratum I population relative abundance of Taxon A were 0.50 instead of 0.40 in the example in Figures 4 and 5, an error would not have occurred because the sample relative abundance given by the Stratum I sampling function would have been greater than the sample relative abundance given

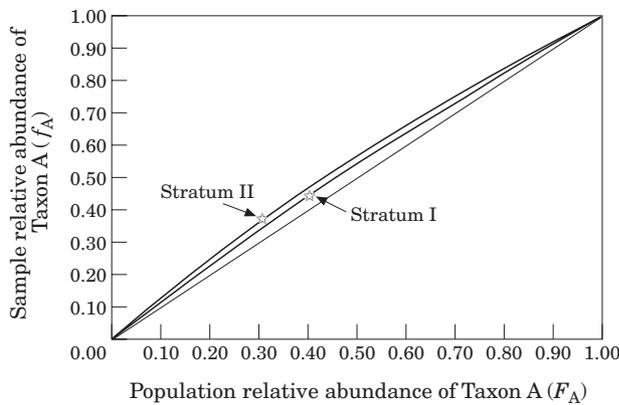


Figure 6. Sampling functions for the 1/8" screen samples in the Figure 4 example.

by the Stratum II sampling function for a Taxon A population relative abundance of 0.30.

The primary implication of this model is that errors in the rank order of samples along an index of relative abundance can only occur when there are differences in the proportion recovered of at least one of the taxa involved between the samples being compared*. For the example in Figure 1, I noted that the rank order of the two strata along the $\Sigma A_i / \Sigma (A_i + B_i)$ index was never incorrect with respect to the rank order of the population values, even though the sample index value of each stratum did vary substantially on an interval level with a change in screen size. Examination of Figure 2 shows why this is the case. If the samples are subject to the same sampling function, then errors in their rank order along an index of relative abundance can never occur because, for any given population relative abundance value, any increase or decrease in population relative abundance will always result in a change in sample relative abundance in the same direction (though not necessarily of the same magnitude).

The converse of the fact that such errors can only occur with differences in sampling functions—i.e., with differences in the average size of the specimens of a given taxon between samples—is that they are much more likely to occur the greater the difference between

* For the sake of simplicity, I have treated the sole effect of increasing fragmentation as being a reduction in the number of specimens recovered. As Marshall & Pilgram (1993) have shown, however, the relationship between fragmentation and NISP is not a linear one. At low levels of fragmentation, increase in fragmentation will cause NISP to increase since the total number of specimens in a sample will increase at a rate higher than the rate at which identifiability decrease. Eventually, though, the decrease in identifiability will overtake the increase in the total number of specimens and NISP will begin to decline at higher levels of fragmentation. Thus, an increase in fragmentation might first cause an increase in the numbers of identifiable specimens recovered in a given-sized screen before it caused a decrease in them. Furthermore, the precise effects of fragmentation are likely to vary greatly between different kinds of taxa and different sizes of screen. For these reasons, it is more appropriate to consider the model presented here as one which deals with differences in the recovered proportions of identifiable specimens, not strictly with differences in fragmentation.

the sampling functions involved. This is where the issue of screen size finally enters. These kinds of errors can occur even when the same size of screen is used to collect multiple samples due to differences between those samples in levels of fragmentation. *If the use of smaller screens results in less biased sampling functions, however, it will reduce the probability that such errors will occur.* This occurs with the 1/8" screen samples in the Figure 4 example, the sampling functions for which are illustrated in Figure 6†. The Stratum II sampling function is still more biased than the Stratum I sampling function with the use of 1/8" screen, but it has become unbiased enough that the rank order of the two samples along the $\Sigma A_i / \Sigma (A_i + B_i)$ index is no longer incorrect. It should also be obvious that any comparison of samples collected with different sized screens, or of screened samples versus non-screened samples, along any measure of relative abundance should be treated with suspicion since the sampling functions involved are likely to be very different.

The Mogollon Village Fauna

Mogollon Village (LA 11568) is a pithouse site located along the San Francisco River in southwestern New Mexico, with dates ranging from approximately AD 120 (and possibly as early as 400 BC) to AD 900 (Mauldin, Gilman & Stevenson, 1996; Linse, 1997; see also Haury, 1936). The faunal samples discussed here were collected in recent excavations by the University of Washington Mogollon Village Field School (MVFS) (Linse, 1997); a detailed analysis of this fauna is presented elsewhere (Cannon, 1999). MVFS screened a random 25% of the sediments from each stratum in each 1 m² excavation unit through 1/8" screen and the rest through 1/4" screen (Hoyer, 1997), and recorded the volume of sediment from each stratum in each excavation unit that was screened through each mesh size (Linse, 1997: appendix D).

Table 1 presents the numbers of identified specimens of artiodactyls and leporids recovered in each size of screen from contexts in and around the two pithouses excavated by MVFS that produced the largest faunal samples, F22 and F43. This table also presents Artiodactyl Index values, calculated after Bayham (1982; Szuter & Bayham, 1989), for these samples. Artiodactyls and leporids are the two most abundant taxa recovered at Mogollon Village, and are also the two taxa that have been considered in previous resource depression analyses using faunas from sites in the

† Figures 4 and 6 also serve to illustrate that the use of a smaller screen must increase the recovered proportion of the less completely recovered taxon by a greater percentage than it increases the recovered proportion of the more completely recovered taxon in order for sampling functions to become less biased. In Stratum I in this example the recovered proportions of both taxa increase by a factor of 3 with the use of 1/8" screen: from 30% to 90% for Taxon A and from 25% to 75% for Taxon B. This results in the Stratum I samples from both screen sizes having the same value along the $\Sigma A_i / \Sigma (A_i + B_i)$ index.

Table 1. Artiodactyl and leporid NISPs and Artiodactyl Index values for the 1/4" and 1/8" screen samples from the F22 and F43 pithouse areas, MVFS fauna

Screen size	Pithouse area	Artiodactyls (NISP)	Leporids (NISP)	$\frac{\Sigma(\text{artiodactyls}_i) \div \Sigma(\text{artiodactyls}_i + \text{leporids}_i)}$
1/4"	F22	25	79	0.240
	F43	52	123	0.297
1/8"	F22	18	62	0.225
	F43	23	124	0.156

southwestern U.S. (e.g., Szuter & Bayham, 1989; Harvey & Broughton, 1996; Janetski, 1997). In this example the taxon Artiodactyla consists of all specimens identified as *Odocoileus* sp. (deer), *Antilocapra americana* (pronghorn), unidentified Cervidae, and unidentified (but deer- and pronghorn-sized) Artiodactyla. The taxon Leporidae consists of all specimens identified as *Sylvilagus* sp. (cottontail rabbits), *Lepus* sp. (hares or "jackrabbits") and unidentified Leporidae.

It can be seen immediately that the 1/4" screen samples from these two pithouse areas show that F43 has a higher abundance of artiodactyls relative to artiodactyls plus leporids than does F22, while the 1/8" samples show just the opposite. If only 1/4" screen had been used here, it appears that an error in the rank order of these two pithouse areas along the Artiodactyl Index would have resulted.

Table 2 shows why this switch in the rank order of these two samples along the Artiodactyl Index has occurred with the change in screen size. The use of 1/8" screen has resulted in a greater percentage increase in artiodactyl specimens recovered per unit volume, and a smaller percentage increase in leporid specimens recovered per unit volume, in F22 than it has in F43. (Since different volumes of sediment were screened through each mesh size by MVFS, the raw numbers of specimens that they recovered in each mesh size are not directly comparable but must be normalized by volume.) Specimens of both artiodactyls and leporids

have responded differently in each pithouse area to the change in screen size.

The "correction factor" method

These empirical findings from the MVFS fauna have serious implications for the "correction factor" method of dealing with screen size biases proposed first by Thomas (1969) and more recently by James (1997). This method involves screening one or more control samples through two or more mesh sizes; from these samples, estimates of the proportions of various taxa recovered in the larger size(s) of screen are calculated. These estimated recovered proportions are then applied as correction factors to samples from provenience units in which only larger screens were used in order to determine what taxonomic abundances would look like in those units if smaller screens had been used. This, of course, assumes that specimens of any taxon are recovered in the same proportion in a given size of screen in all of the provenience units that are being compared. The Mogollon Village fauna shows that such an assumption is unwarranted: the fact that both artiodactyls and leporids have responded differently between pithouse areas to a change in screen size strongly implies that the proportions of these taxa recovered in a given screen size do vary between the two provenience units considered here (see Lyman, 1992 for a similar point). This is most likely the result of variation in the average size of the specimens of each taxon between the two pithouse areas.

It can also be shown that the correction factor method produces values that are much different from values actually observed. Table 3 presents artiodactyl and leporid abundances for the F22 and F43 areas "corrected" according to the method used by James (1997). James's correction factors are calculated as $(1/4" \text{ NISP}_i + 1/8" \text{ NISP}_i) / 1/4" \text{ NISP}_i$ (James 1997: table 4). The correction factors used here are calculated in this manner, with the exception that I have used the volume-normalized NISP values of Table 2 in order to account for the fact that different volumes of sediment were screened through each mesh size by MVFS. The "corrected" artiodactyl and leporid values for each

Table 2. Numbers of identified artiodactyl and leporid specimens recovered per unit volume of excavated sediment in the 1/4" and 1/8" screen samples from the F22 and F43 pithouse areas, MVFS fauna

Pithouse area	Screen size	Excavated sediment volume (m ³)	Artiodactyls (NISP)/volume(m ³)	Percentage increase in artiodactyls/m ³ with 1/8" screen	Leporids (NISP)/volume(m ³)	Percentage increase in leporids/m ³ with 1/8" screen
F22	1/4"	5.14	4.9		15.4	
	1/8"	1.61	11.2	229%	38.5	250%
F43	1/4"	8.12	6.4		15.1	
	1/8"	2.53	9.1	142%	49.0	325%

Table 3. Comparison of observed values and values resulting from "correction" according to the method proposed by James (1997) for the MVFS fauna

Sample	Artiodactyls (NISP)/volume (m ³)	Artiodactyl correction factor	Leporids (NISP)/volume (m ³)	Leporid correction factor	$\Sigma(\text{artiodactyls}_i \div \Sigma(\text{artiodactyls}_i + \text{leporids}_i))^a$
F22 1/4" observed	4.9		15.4		0.241
F43 1/4" observed	6.4		15.1		0.298
F22 1/8" observed	11.2	3.3	38.5	3.5	0.225
F43 1/8" observed	9.1	2.4	49.0	4.2	0.157
F22 1/8" "corrected"	11.8		64.7		0.154
F43 1/8" "corrected"	21.1		52.5		0.287

^a The slight differences between some of the observed Artiodactyl Index values here and those in Table 1 are the result of rounding errors that occurred in the volume-normalization process.

The "corrected" 1/8" values for each pithouse area are calculated by multiplying the observed 1/4" values for that pithouse area by the correction factors derived from the other. For example, the "corrected" artiodactyl/m³ value for F22 (11.8) is the observed F22 1/4" value (4.9) multiplied by the F43 artiodactyl correction factor (2.4).

pithouse area are determined by multiplying the observed 1/4" values for that pithouse area by the correction factors derived from the other pithouse area; Artiodactyl Index values derived from these "corrected" NISPs are also presented. It is obvious that the "corrected" values for each pithouse area are much different from the observed 1/8" screen values no matter which of the two areas is used as the control sample. Furthermore, comparison of the "corrected" Artiodactyl Index value for either pithouse area with the observed 1/8" screen value for the other shows that this "correction" merely preserves the rank order of the two pithouse areas along the Artiodactyl Index that is observed in the 1/4" screen samples. By assuming that recovered proportions are always going to be the same between provenience units, the correction factor method precludes one from addressing the possibility that they might be different.

Discussion

The differential recovery model and the empirical findings from Mogollon Village are directly relevant to zooarchaeological analyses that are based on differences between samples in the abundance of a large-bodied taxon relative to a small one; the resource depression studies discussed above are a prime example of this kind of analysis. The main implication of the differential recovery model for these studies is that they will *not* be subject to screen size-induced errors if each taxon involved is recovered in an equivalent proportion in each of the samples being used. In other words, if the degree of fragmentation or the average specimen size of each taxon is the same between provenience units, then the patterns found in such studies will be robust with respect to the differential recovery of taxa that vary in body size. This will be the case no matter what size of mesh was used in excava-

tion, or even if no screens were used, provided that the same collection protocol was used for all samples.

The Mogollon Village fauna, though, suggests that differences in recovered proportions between samples might often exist, even between samples taken from very similar depositional contexts. Not only does the possibility that such differences might exist render the use of "correction factors" problematic, it also poses a problem for analyses that rely on temporal or spatial trends in relative abundance indices. If differences in recovered proportions between provenience units are present in such an analysis, it is possible that any trends that are evident may simply reflect differential recovery rather than actual population differences. Analysis of the degree of fragmentation per taxon per provenience unit might be useful for assessing independently whether this is likely to be a problem in a given study (see Grayson & Delpech, 1998 for a similar point in the context of diet breadth analysis).

Of course, if differences between provenience units in degrees of fragmentation or average specimen size per taxon are present in a given case, this does not necessarily mean that differential recovery errors will occur. The probability of this kind of error occurring will increase, however, with the use of larger sizes of mesh in excavation. Analyses of archaeofaunas excavated prior to the routine use of screens (e.g., Broughton, 1994a, b, 1995, 1997; Szuter & Bayham, 1989) will thus be particularly susceptible, and the extent to which differences in recovered proportions between provenience units are present in these analyses should be of primary concern.

Conclusions

This presentation is not intended to argue that smaller screens are "better", a point which has been made countless times before. Rather than just making the

generalization that smaller screens are “better”, the differential recovery model allows us to see when they will be better and why. The model also shows how variability in certain factors that might not have been considered before, such as variability in population relative abundances or variability in degrees of fragmentation between samples, can influence the probability that differential recovery errors might occur. Combined with the empirical findings from Mogollon Village, the model suggests that fragmentation levels per provenience unit should be assessed independently in analyses of spatial or temporal trends in the abundance of one taxon relative to another. The Mogollon Village fauna also calls into question the utility of “correction factor” methods for dealing with screen size bias, since it is possible that a given taxon can respond differently to a change in screen size in different provenience units.

Archaeologists have long recognized that the choice of mesh size made in an excavation can impact the nature of the faunal data that it produces. The Mogollon Village fauna suggests that these impacts can be substantial and that errors in the results of analyses due to differential recovery might be common when larger mesh screens are used. Rather than just reiterating that data may be biased as the result of screen size choice, however, the model presented here allows one to determine the conditions under which specific kinds of errors will be more or less likely to occur and to see exactly how screen size choice can affect the probability that they will occur. The accurate quantification of faunal data is crucial to understanding prehistoric subsistence, but truly accurate quantification requires an understanding of, and an ability to deal with, potential biases in this kind of data.

Acknowledgements

I thank Donald K. Grayson, Jack M. Broughton, R. Lee Lyman and two anonymous reviewers for very helpful comments on earlier drafts of this paper. The University of Washington Mogollon Village Field School was supported by the University of Washington and the USDA Forest Service, Gila National Forest, Silver City, NM; I thank Angela R. Linse for providing access to the fauna from this project. My analysis of the Mogollon Village fauna was aided by a travel grant from the Department of Anthropology, University of Washington.

References

- Bayham, F. E. (1979). Factors influencing the archaic pattern of animal utilization. *Kiva* **44**, 219–235.
- Bayham, F. E. (1982). *A Diachronic Analysis of Prehistoric Animal Exploitation at Ventana Cave*. Ph.D. Thesis. Arizona State University.
- Broughton, J. M. (1994a). Declines in mammalian foraging efficiency during the Late Holocene, San Francisco Bay, California. *Journal of Anthropological Archaeology* **13**, 371–401.
- Broughton, J. M. (1994b). Late Holocene resource intensification in the Sacramento Valley, California: the vertebrate evidence. *Journal of Archaeological Science* **21**, 501–514.
- Broughton, J. M. (1995). *Resource Depression and Intensification during the Late Holocene, San Francisco Bay: Evidence from the Emeryville Shellmound Vertebrate Fauna*. Ph.D. Thesis. University of Washington.
- Broughton, J. M. (1997). Widening diet breadth, declining foraging efficiency, and prehistoric harvest pressure: ichthyofaunal evidence from the Emeryville Shellmound, California. *Antiquity* **71**, 845–862.
- Butler, V. L. (1993). Natural versus cultural salmonid remains: origin of the Dalles Roadcut bones. *Journal of Archaeological Science* **20**, 1–24.
- Cannon, M. D. (1999). The faunal remains from Mogollon Village (LA 11568). To be submitted to U.S.D.A. Forest Service, Gila National Forest, Silver City, New Mexico.
- Casteel, R. W. (1972). Some biases in the recovery of archaeological faunal remains. *Proceedings of the Prehistoric Society* **38**, 383–388.
- Casteel, R. W. (1976). Comparison of column and whole unit samples for recovering fish remains. *World Archaeology* **8**, 192–196.
- Charnov, E. L., Orians, G. H. & Hyatt, K. (1976). Ecological implications of resource depression. *The American Naturalist* **110**, 247–259.
- Grayson, D. K. (1984). *Quantitative Zooarchaeology*. New York: Academic Press.
- Grayson, D. K. (1991). Alpine faunas from the White Mountains, California: adaptive change in the prehistoric Great Basin? *Journal of Archaeological Science* **18**, 483–506.
- Grayson, D. K. & Cannon, M. D. (1999). Human paleoecology and foraging theory in the Great Basin. In (C. Beck, Ed.) *Models for the Millennium: The Current Status of Great Basin Anthropological Research*. Salt Lake City: University of Utah Press (in press).
- Grayson, D. K. & Delpech, F. (1998). Changing diet breadth in the early Upper Palaeolithic of southwestern France. *Journal of Archaeological Science* **25**, 1119–1130.
- Haurly, P. K. & Broughton, J. M. (1996). Variation in hunting efficiency during the Fremont: archaeological vertebrate evidence from the Evans Mound, southwestern Utah. Paper presented at the 25th Great Basin Anthropological Conference, King's Beach, California.
- Haurly, E. W. (1936). *The Mogollon Culture of Southwestern New Mexico*. Medallion Papers **20**. Gila Pueblo, Globe, Arizona.
- Hildebrandt, W. R. & Jones, T. L. (1992). Evolution of marine mammal hunting: a view from the California and Oregon coasts. *Journal of Anthropological Archaeology* **11**, 360–401.
- Hoyer, M. E. (1997). Excavation strategy. In (A. R. Linse, Ed.) *Excavations at Mogollon Village and survey of surrounding areas: University of Washington 1993 Archaeological Field School*. Submitted to U.S.D.A. Forest Service, Gila National Forest, Silver City, New Mexico, pp. 10–11.
- James, S. R. (1997). Methodological issues concerning screen size recovery rates and their effects on archaeofaunal interpretations. *Journal of Archaeological Science* **24**, 385–397.
- Janetski, J. C. (1997). Fremont hunting and resource intensification in the eastern Great Basin. *Journal of Archaeological Science* **24**, 1075–1088.
- Linse, A. R. (1997). Excavations at Mogollon Village and survey of surrounding areas: University of Washington 1993 Archaeological Field School. Submitted to U.S.D.A. Forest Service, Gila National Forest, Silver City, New Mexico.
- Lyman, R. L. (1992). Review of A. Cannon, 1991, *The Economic Prehistory of Namu*. *Canadian Journal of Archaeology* **16**, 134–136.
- Marshall, F. & Pilgram, T. (1993). NISP vs. MNI in quantification of body-part representation. *American Antiquity* **58**, 261–269.
- Mauldin, R., Gilman, P. A. & Stevenson, C. M. (1996). Mogollon Village revisited: recent chronometric results and interpretations. *Kiva* **61**, 385–400.
- Nagaoka, L. (1994). Differential recovery of Pacific island fish remains: evidence from the Moturakau Rockshelter, Aitutaki, Cook Islands. *Asian Perspectives* **33**, 1–17.

- Payne, S. (1972). Partial recovery and sample bias: the results of some sieving experiments. In (E. S. Higgs, Ed.) *Papers in Economic Prehistory*. Cambridge: Cambridge University Press, pp. 49–64.
- Shaffer, B. S. (1992). Quarter-inch screening: understanding biases in recovery of vertebrate faunal remains. *American Antiquity* **57**, 129–136.
- Shaffer, B. S. & Sanchez, J. L. J. (1994). Comparison of 1/8"- and 1/4"-mesh recovery of controlled samples of small-to-medium-sized mammals. *American Antiquity* **57**, 525–530.
- Stahl, P. W. (1996). The recovery and interpretation of micro-vertebrate bone assemblages from archaeological contexts. *Journal of Archaeological Method and Theory* **3**, 31–75.
- Szuter, C. R. & Bayham, F. E. (1989). Sedentism and prehistoric animal procurement among desert horticulturalists of the North American Southwest. In (S. Kent, Ed.) *Farmers as Hunters: The Implications of Sedentism*. Cambridge: Cambridge University Press, pp. 80–95.
- Thomas, D. H. (1969). Great Basin hunting patterns: a quantitative method for treating faunal remains. *American Antiquity* **34**, 393–401.