<III. Methodology>

<Introduction>

We presented our orientation and some of our specific hypotheses in the previous chapter and in this chapter describe the procedures we used to obtain observations relevant to those ideas and how we will manipulate the observations to examine our ideas about Cedar Mesa adaptations. Our methodology can be divided into three sections: sampling procedures, field procedures and analytical procedures.

The sampling procedures described below include some important discussions about assumptions involving the past environment. Such decisions as how the mesa was divided up, how quadrats were selected and how sites in the drainage canyons were sampled are discussed in the light of sampling theory, and of our knowledge about the environment and archaeology. Other criteria were used to decide which sites were sampled in the architectural tree-ring study which will be fully reported in the future. The same field procedures were used for collecting sites found in quadrats and in canyons of the five sampled watersheds. These procedures are described below. In addition to this set of procedures, a separate one was used to inventory the drainage canyon sites, prior to sampling them.

Data analysis proceeded in two fashions: 1) by direct
Matson, Lipe, and Haase (Aug. 88) III-2 tests of specific hypotheses, either developed prior to the start of the project or during it, and 2) data reduction to group the large number of sites and descriptive variables into more manageable units. The direct hypothesis testing was usually carried out by non-parametric analysis of variance and the data reduction by clustering and multi-dimensional scaling (True and Matson 1970, Matson and True 1974, Matson 1974). There is not a temporal or logical order between these two procedures; as analysis of variance was carried out prior to data reduction, as well as after it, using the results of clustering and scaling. The suitability of these techniques and of the assumptions made in order to apply them are discussed below.

<Sampling Design>

We have argued above for the <regional> approach, where a relatively large area is the subject of investigation and where the main research interest is on questions of environment associations, settlement pattern, and settlement system. Because Cedar Mesa was too large to be completely surveyed, we had to use a sampling scheme. Questions which had to be resolved included whether there were important trends in Cedar Mesa geography which would call for stratification, whether natural units existed that should be subsampled, and what size and shape the actual surveyed units should be. We decided to stratify the mesa into North and South halves and to treat the canyon bottoms and mesa tops differently. Natural units were seen to exist in
Matson, Lipe, and Haase (Aug. 88) III-3 drainages which were used to subdivide the study area; and the actual survey units selected were quadrats 400 meters on a side. In addition, the sites in the canyon bottoms were later inventoried, and a sample of these sites were selected for further study.

<Drainage Units>

One important decision was to treat Cedar Mesa as consisting of drainage units. We perceived drainage units as legitimate subdivisions of Cedar Mesa because each one included the main types of resources we thought were important to prehistoric Anasazi inhabitants. One important resource type was the springs, which typically occur in the canyons and at the points where the drainages begin to entrench (Figure III-1). If a nearby source of water was important (and we will show later that for some site types and times it was), habitation sites should be located with respect to springs and, therefore, to drainages. The deep soil, sagebrush flats usually occur on the mesa-tops in broad shallow valleys that lie between the central north-south mesa divide and the points where the drainages become entrenched in canyons. If sagebrush flats were important farming areas, (and we will later examine this notion thoroughly) sites should be located with respect to them.

Earlier, we argued that deep soil areas on the divides were important agricultural locations for the prehistoric inhabitants. Within each drainage unit, such deep soil
Matson, Lipe, and Haase (Aug. 88) III-4
divide areas exist, both along the mesa’s north-south divide
and along the east-west trending ridges which are the
borders between drainage units. Also present in each
drainage unit are thin soiled, slickrock canyon rims, and
the entrenched canyons themselves. The latter are the most
biotically diverse environments in the study area, as
indicated earlier. Further, within each drainage unit, the
range of elevations is usually a substantial portion of that
found on the study area as a whole. In short, much of the
environmental variation that exists on Cedar Mesa is found
within each drainage. Drainage units are obvious
environmental subdivisions on Cedar Mesa and are to be
preferred over arbitrary clusters.

An advantage of dividing Cedar Mesa into some set of
natural "clusters" is that this would allow us to make
statements and comparisons to be made about natural
subareas, which would be difficult to do in their absence.
The disadvantages of using such natural units are that the
actual boundaries are often arbitrary and that large
differences in areal size may occur. In this case, the
sizes did not differ too radically and by subdividing a few
of the larger ones we split the mesa into twenty drainages
of roughly equal size.

When a population is divided into a series of units that
are larger than the population elements (quadrats in this
case), the selection of such units is said to be cluster
sampling (Cochran 1963:235). So we made the decision to use
drainages as the first order of sampling units.
An alternative to the selection of clusters would have been to draw a simple random sample of quadrats from Cedar Mesa at large. Surely such an approach would have resulted in samples of the various environmental zones of interest. Using our time and budget, this approach would have allowed between 3 and 4 quadrats per drainage, too small to compare drainages or to discover intradrainage settlement patterns. Thus to compare portions of the mesa, areas of at least two to three drainages in size would have to be lumped together, at best, and details about drainages would be obscured.

Finally our assumption that drainages were natural units would not be exploited. While a simple random sample might be an efficient way of making generalizations about the study area as a whole, it would not give good information about local differences.

Since cost efficiency is a central aspect of sampling (an inexpensive way to make inferences about the whole), the relative costs should also be examined. With the simple random approach, because the quadrats would be more widely dispersed, more time would be spent in finding quadrats than with the cluster scheme used. Thus in terms of amount of time surveying versus amount spent on logistics, the simple random approach would be less efficient. The major advantage that the simpler approach does have is in its simplicity, which as we will see later, is no small advantage.

<Stratification of the Drainages>
From our previous work on Cedar Mesa, we knew environmental differences between the northeast and southwest portions existed. While the mesa as a whole is dominated by pinyon-juniper and sagebrush flats, these communities do vary in different places. In general the amount of pinyon-juniper increases as one goes north, as well as the ratio of pinyon to juniper, which we reported in Chapter 1. The relative importance of sagebrush flats, conversely increases as one goes south. The blackbrush community is found primarily in the southwest. Besides these vegetation differences, a physiographic variable also varies from north to south; areas of loose aeolian and probably recently reworked sediments are much more common to the south. Both the vegetation and physiographic differences may be related to the precipitation pattern discussed previously, with increased rainfall in the north.

Since the known environmental trends are along a north-south trend, if the cluster samples of drainages were to be stratified it would be on this basis. We also had impressions from sketchy archaeological reconnaissance that archaeological differences existed on a north-south(or northeast-southwest) trend as well. Further work (see Chapter 4) confirmed that strong north-south trends in the archaeological material existed, but not the expected Kayenta style Pueblo II ceramics in the south and Pueblo II-III Mesa Verde ceramics in the north. Because both the environmental trends and suspected archaeological trends were on a north-south basis, we chose to stratify the
Matson, Lipe, and Haase (Aug. 88) III-7

drainages in that manner, mainly to ensure that areas from
both ends of the continuum were sampled to allow these
differences to be tested and inspected in detail. Because
we could afford to sample only a few drainages we feared
that unless we stratified the sample, all our samples might
fall into either the north or the south.

A relatively arbitrary line was picked to separate the
drainages into north and south strata. From this point 12

and 8 to the southwest.

Sampling the drainages at a 25 percent rate would mean the
selection of two from the south and three from the north.
The 25 percent figure was picked on the basis of average
size of drainage units, expected budget and expected cost
per quadrat to give a reasonable sample per drainage.

Our decision to use natural drainage units instead of
artificial ones as clusters posed the problem of unequal
size. This was alleviated by assigning the probability of
being chosen to a drainage units in proportion to its size.
By making the probability of being selected proportional to
the amount of area within a drainage any given area has an
equal probability of being sampled (Cochran 1963:251). Thus
even with clusters of unequal sizes, the basic concept of
equal probability sampling (and simpler formulas) can be
used (Cochran 1963:297). We followed this procedure to
select two clusters, West Johns and Hardscrabble, from the
south and three, Upper Grand Gulch, Bullet, and North Road,
from the north (Figure III-2).

Other ways of selecting drainages may have been as
Matson, Lipe, and Haase (Aug. 88) III-8 appropriate. For instance, some sort of systematic sampling, such as randomizing a beginning point and then picking every fourth drainage, might have been as unbiased and more straightforward. This would not be equal probability sampling, as areas in smaller drainages would have a higher probability of being selected than those in larger drainages. Another possible drawback would be the chance of hidden periodicity in the drainages. On the first glance such a patterning in the data would seem unlikely. On the other hand, Hudson’s theory of colonization (Hudson 1969) suggests that in a developed settlement system, competition between centers develops and that this leads to relatively equal spacing. Certainly there is some evidence of defensive structures in late Pueblo III times that might be interpreted as support for the notion of competition between centers. This sort of situation might create the kind of "periodicity" that makes problems for a systematic sample (Cochran 1963:230). While both of these concerns may be marginal, they do illustrate that there are benefits to using the stratified random sampling approach, even though it may be more complicated.

<Second Stage Sample Units: Quadrats>

Although the area enclosed by the five selected drainages is small relative to Cedar Mesa as a whole, it is still much too large to completely survey in an intensive fashion. Thus we had to subsample each drainage, or use a two stage design. A series of decisions were made concerning the kind of unit to be used and whether this
Matson, Lipe, and Haase (Aug. 88) III-9
subsample should itself be stratified.

There are two choices for the shape of the sampling
unit—square quadrat or narrow transect. While Binford
(1964) in his pioneering paper on research design suggested
quadrats and initial efforts to apply sampling theory to
regional survey quadrats (Matson 1971, 1974; Thomas 1970;
Plog 1974), some more recent work has favored transects
(Plog 1976; Judge, Ebert and Hitchcock 1975). The last
report, unfortunately, apparently ended up comparing the
population with the population for estimating the reliability
of systematic transects (1975:116;115). Because of the
non-random nature of systematic samples, it is difficult to
calculate satisfactory estimates of precision, but clearly
that procedure is invalid. The random transects reported on
in Judge, Ebert and Hitchcock, however, do not share this
problem and compared favorably with quadrats. In a very
extensive simulation, Plog found that the most consistent
gains in precision resulted from a decrease in size of
sample units, and secondly, in the use of transect samples
rather than quadrat samples (1976:157). Both of the above
studies used estimates of total populations as measures of
relative efficiency. In plant ecology, the transect has
long been shown to be the most efficient shape for obtaining
population parameter estimates (Daubenmire 1968:87-88). The
reason transects are more efficient than quadrats is,
however, a drawback for our situation.

To quote Daubenmire "Since the individuals representing
each species tend to be grouped into isodimetric clusters,
an elongate plot has a high probability of intercepting
parts of several clusters at once without falling entirely
into one" (Daubenmire 1968:88). This statement might be
translated into archaeological terms as "since a transect
plot will crosscut more sites and environmental diversity
than a quadrat of equal size, it thus will be more
representative of an area." Since, however, we are not only
interested in population estimates but also in associations,
we would prefer to minimize the environmental change within
each unit and maximize the area surrounding each site within
each sampling unit. Another disadvantage with a transect is
that it has more margin per unit area and thus the problem
of sites located on unit edges is magnified, as compared to
quadrats. For these reasons, we chose to use quadrats
instead of transects.

Another important factor is the size of the sampling
unit. One aspect of this factor is the "grab" sample
effect. That is, items found in any continuous unit tend to
be more like each other than items randomly chosen from the
population at large. An extreme example would be to
contrast the representativeness of a sample consisting of
one quadrat 1600 meters on a side with that of 16 randomly
chosen quadrats 400 meters on a side, two samples of equal
area. Obviously the sample of 16 will be more spread out
over the population and be more representative than the
equal sized (in area) sample of one large quadrat. Thus the
representativeness of a sample increases as the sample units
become smaller in size and larger in number, everything else
Matson, Lipe, and Haase (Aug. 88) III-11 being equal. In sampling simulations carried out prior to the design of this project this factor consistently shows up (True and Matson 1974, Plog 1976) in spite of a more recent comment on its supposed non-existence (B. Laird Hole 1980).

Size is also a factor in the compatibility of sampling unit and stratum. If the strata are small and the quadrats are large, error will occur through quadrats overlapping two or more environmental strata. Since all of the statistics used are more dependent on the size of the sample as opposed to the sampling rate, the larger the quadrats, the larger the percentage of the stratum area that will have to be sampled in order to obtain a sufficient number of sampling units in order to make significant comparisons.

Although the above effects favor decreasing the size of the quadrats, there are others, primarily economic ones, which go the other way. For surveying any given area, decreasing the quadrat size also decreases the proportion of time spent surveying and increases the proportion spent getting to the quadrat, locating and marking its boundaries, etc. Another problem also increases as the size decreases, that of sites lying on sample unit borders. This can be a very serious problem with transects, as mentioned above, and also occurs with quadrats. Quadrats must be large enough to completely include, on the average, the sites being surveyed. Otherwise to get complete information on sites one would have to spend more time outside the quadrat than inside. S. Plog (1976) found that with transects it was necessary to calculate how much of the site was inside the
Matson, Lipe, and Haase (Aug. 88) III-12 sample unit border to calculate actual population estimates and we expect that this would also be necessary with small quadrats. As quadrats decrease in size the ratio of edge to area increases, making this problem worse (Matson, Magne, Ludowicz and Pokotylo 1980:31).

The final, and probably most important, factor in quadrat size is that quadrats need to be sufficiently large so that they usually include some archaeological material. While some empty quadrats can be tolerated, they should be in the minority; otherwise methods of analysis become difficult to use, and even descriptive statistics become less informative because of the extensive quadrats tied at zero.

The usefulness of any quadrat size also depends on the distribution of the variables involved. Since in our project, a large number of variables are of interest, the best size must, by definition, be a compromise. Further, we were more interested in obtaining useful samples rather than in most efficiently calculating population parameters of one single variable, the usual criterion used in sampling evaluations. In any case, since the distributions of the variables of interest were unknown, it was impossible to calculate precisely the best sample unit size for estimating population parameters for any or all variables.

In the seminal work which influenced so much regional survey work in the 1970’s, Binford (1964) suggested the use of quadrats one-half mile on a side. The first applications used quadrats 500 meters on a side (Matson 1971,1974, Thomas
Matson, Lipe, and Haase (Aug. 88) III-13 1969). Matson and Thomas found sufficient information for their projects using 500 meter quadrats. Since Cedar Mesa site and artifact densities are much higher than those in the Cerbat Mountains, Arizona, and Reese River, Nevada, where Matson and Thomas worked, a smaller quadrat size was feasible without generating excessive numbers of empty quadrats. As pointed out above, most analysis of variance type statistics are dependent on the number of sampling units, as opposed to size of sample units(or sampling rates), so that for a given length of time in the field, reducing the size of the quadrats will increase the sensitivity of the sample. Taking all these factors into account, we decided on square quadrats 400 meters on a side. Since the size of our five selected drainages range from 20 to 51 square kilometers( 8 to 19 sq. mi.) the number of 400 meter quadrats in each drainage ranges from 126 to 308 . In the actual field work we sampled at a 7% rate which resulted in samples that ranged from 9 to 22 with a total of 76 quadrats.

<Quadrat Sampling Method>

Once the size and shape of the sampling units is established, the method of selection yet remains. Again referring to the botanical literature we find that:

Systematic sampling...is ...more advantageous if interest centres on variability within the area....Of the two it is clear that, in most circumstances random sampling is the better if overall information on the composition is desired (Greig-Smith 1964:23).
It would appear, then that random sampling is most appropriate for our purposes. Cochran, in a discussion of systematic sampling asserts it can be safely recommended under four conditions.

The first is where the ordering of the population is essentially random, the second where a stratified design with numerous strata is used, the third is for subsampling the unit, and the fourth is for sampling populations of a continuous type (Cochran 1963:230)

The first two cases are clearly not the situation on Cedar Mesa. The last situation is also inappropriate, as archaeological variables occur in discrete rather than continuous fashion. The third condition at first appears to be similar to our situation with drainages and quadrats, but a close inspection of the examples indicates some significant differences.

Because of the irregular sizes and shapes of the drainages, it is difficult to know how to set up a systematic sample. If one uses some kind of geographical pattern of systematic pattern, such as selecting the center quadrat of a certain size repeated grid, one is likely to end up with a sample heavily weighted in divide areas in one drainage, and rim areas in another. These results would occur because of the different size of the drainages and thus the different fits of the grids within the different drainages. It does not seem reasonable to use different patterns in different drainages to try to avoid this. The
Matson, Lipe, and Haase (Aug. 88) III-15 other systematic pattern of numbering the quadrats in a linear order and selecting every 14th (approximately 7%) seems to be a perversion of the meaning of a systematic samples, as this would result in very different patterns if one used the same method for all drainages, again due to their different sizes and shapes. Using different ways of ordering the lists for different drainages does not seem to be any more appropriate than using different patterns. It is probable that the subsampling referred to by Cochran involves first stage units of regular size (and shape if translated into geographical terms).

In addition to the non-archaeological reasons for dismissing systematic samples, S. Plog's empirical study found systematic samples gave highly variable results (1976:150) and little gain in precision. For these reasons we eliminated systematic sampling from further consideration, although we are left with doubts that choosing any of these really have made much difference in the end from the random sampling scheme we used.

Another approach which has been suggested is the systematic-unaligned sampling scheme in which one coordinate remains fixed and one random for each interval row (Berry and Baker 1968; Redman and Watson 1970). This procedure was also tested by Plog and was found to be the least efficient of the designs studied (Plog 1976). It should be noted that this design was shown to be superior to random sampling (Berry and Baker 1968) in a study of land use or land type; this variable has nearly continuous variation while sites or
Matson, Lipe, and Haase (Aug. 88) III-16 artifacts do not. Again this type of sampling has both theoretical and pragmatic disadvantages that seemed to make it a poor choice for our purposes. Both of the above alternatives to random sampling have a further limitation in that the estimate of sampling error is either unobtainable or of questionable validity (Cochran 1963:230). It seems that these approaches should not be used under other situations than those mentioned above. Only when those conditions are met do these methods have advantages that outweigh the inability to estimate sampling error. In sum, some form of random sampling appeared to be the most appropriate choice of sampling the drainages with quadrats. The question of further stratifying the area within the drainages also had to be answered. A stratified design would be justified on the basis of known or suspected homogeneity of some portion of the drainages. We did expect that some environmental zones would be more homogenous than the drainages as a whole. The most obvious difference is the entrenched canyons versus the mesa top. Within the mesa top possible subdivisions included areas along the main divides versus the rest of the mesa top and sagebrush dominated vegetation versus pinyon-juniper dominated vegetation. The stability of the vegetation contrasts is, however, undemonstrated and the divide versus non divide dichotomy is difficult to deal with, being a continuous, rather than a discrete variation, as we will demonstrate during the analysis. Of the possible subdivisions, we decided to recognize only the canyon versus
Matson, Lipe, and Haase (Aug. 88) III-17
the mesa-top as separate strata. The other physiographic
contrasts are valid and the reader will note numerous
references to them later. We incorporated them by post
stratifying the sample. By this we mean grouping the
sampled elements into strata, such as divide and non-divide.
This after-the-fact stratification has the advantage that
much more information will be available on environmental
variables after the survey than prior to it; the danger is
that the sample of archaeological associations with
particular variables may not be large enough. In addition,
comparisons between drainages may not be possible for
certain strata, because they were not built into the
sampling design. We did not think that these factors
warranted the extra complications and problems involved. In
fact some of the post stratification we have used is based
on environmental information collected during the survey,
but which was not available prior to the fieldwork.

<Sampling the Drainage Canyons>

Because the canyon versus mesa-top is the greatest
environmental contrast in all five drainages we decided to
use this contrast for stratification. We had initially
planned to use double sampling, that is, to sample the
entire drainage at a constant rate and then resample the
canyon strata with additional quadrats to give it a higher
sampling rate (Matson and Lipe 1975:134). After completing
the basic 7 percent quadrat survey we did not think we had
an adequate sample from the canyons but recognized problems
Matson, Lipe, and Haase (Aug. 88) III-18 with the proposed approach. The actual area of the canyons was small, ranging from 6 to 28 percent of the drainage area. Four of the five drainages have between 6 and 15 percent of their area in canyons, but with West Johns has a much higher percentage, in part because of how it was defined, and in part because of its position at the edge of the mesa. Using different physiographic definitions of canyon would result in somewhat different values.

To get a sufficient sample size of quadrats from each canyon would mean sampling at a high rate, which is usually an inefficient way of operation. The canyons themselves are usually narrow and twisting making 400 meter quadrats difficult to use (see our comments above about quadrat and strata size). Using a different quadrat size would make the boundary between canyon and mesa top strata very difficult to work with, as a common boundary would be difficult to develop because of the differing quadrat sizes. Problems in analysis also would occur because of different quadrat sizes. In addition, it became clear that quadrats hanging over canyon walls become highly irregular, and are difficult and dangerous to survey as well as uneconomical. Still, we thought that more information about the canyons, particularly during late Pueblo III times, was highly desirable.

The alternative we selected was quite different from the previously discussed quadrat procedures. We decided to develop a list of drainage canyon sites and sample from this list for detailed artifact collection, recording, and
Matson, Lipe, and Haase (Aug. 88) III-19 mapping. In the quadrat approach we are sampling chunks of space with the archaeological material being aspects of that space; in the drainage canyons we were sampling the sites directly as elements. The list was developed by first doing an inventory of all drainage canyon sites. A complete inventory of the five drainage canyons was only possible because of the relatively small area involved. The inventory allowed us to stratify the sites by types developed on the basis of the inventory. We sampled the different sites types at different rates, treating each drainage separately.

The eventual scheme developed was one focussed on obtaining useful samples rather than statistical elegance. On the basis of the inventory forms six site classes were developed: 1. Basketmaker II.

2. Basketmaker III Any site having a Basketmaker III component was placed in this class.

3. Pueblo Habitation Sites.

4. Pueblo Non-habitation Sites. These were predominantly graneries.

5. Specials. This class consists of sites such as defensive works and isolated kivas that were of interest and/or unusual.

6. Others. None of the above: historic sites, pictograph panels, unclassified sites.

For the Basketmaker II sites a sample size of seven was selected, so that if a drainage had less than 8 Basketmaker II sites, all were collected; if more existed, a random
Matson, Lipe, and Haase (Aug. 88) III-20 sample of seven was taken.

For Basketmaker III sites, which were defined to include multicomponent sites, the same procedure was used, only no drainage canyon had more than seven Basketmaker III sites (Table III-1).

Five sites of both habitation and non-habitation Pueblo categories were selected. This size sample was found in all drainages as abundant Pueblo sites of all kinds existed.

All the "Specials" and none of the sites in the residual category were collected.

By these procedures we attempted to develop samples of canyon sites that could usefully be compared with mesa top sites.

The actual sample of drainage canyon sites was larger than would be expected from the above. First, some quadrats did fall into the canyons and had sites in them. Second, on the first sampling, all the Pueblo non-habitation sites selected for the North Road drainage were in the upper half of the drainage, so an additional three sites in the lower half were selected later.

This sampling scheme was developed only after two years of mesa-top survey and considerable preliminary analysis of the data in addition to the inventory of two drainage canyons and half of a third. It was carried out as a supplement to the main quadrat survey. When the selected canyon sites were revisited, it was found that the inventory notes were not always precise (See note at the bottom of Table III-1). In some cases, if the sites sampled turned
Matson, Lipe, and Haase (Aug. 88) III-21 out to be very small, additional members of a class were selected.

Table III-1

<Drainage Canyon Sampling Summary>
(Not including quadrat canyon sites)

<table>
<thead>
<tr>
<th>&lt;Drainage</th>
<th>Site Class</th>
<th>Inventoried</th>
<th>Sampled</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM II</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>BM III</td>
<td>3</td>
<td>4*</td>
<td></td>
</tr>
<tr>
<td>Upper</td>
<td>PII/III Habitation</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Grand</td>
<td>PII/III Non-Habitation</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>Gulch</td>
<td>Specials</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&lt; Other</td>
<td>0</td>
<td>0&gt;</td>
<td></td>
</tr>
</tbody>
</table>

| Bullet    | PII/III Habitation      | 26          | 5       |
|           | PII/III Non-Habitation  | 31          | 5       |
|           | Specials               | 5           | 5       |
| < Other   | 5                       | 0>          |

<p>| North     | PII/III Habitation      | 10          | 5       |
| Road      | PII/III Non-Habitation  | 25          | 8       |
|           | Specials               | 1           | 1       |
| &lt; Other   | 2                       | 0&gt;          |</p>
<table>
<thead>
<tr>
<th></th>
<th>BM II</th>
<th>BM III</th>
<th>PII/III Habitation</th>
<th>PII/III Non-Habitation</th>
<th>Special</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard-scrabble</td>
<td></td>
<td></td>
<td>5</td>
<td>8</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

|                     |       |        |                    |                       |         |       |
| BM II               | 25    | 4      |                    |                       |         |       |
| BM III              | 1     | 1      |                    |                       |         |       |
| West                |       |        |                    |                       |         |       |
| PII/III Habitation  | 17    | 5      |                    |                       |         |       |
| Johns               |       |        |                    |                       |         |       |
| PII/III Non-Habitation | 24 | 5 | |                       |         |       |
| Special             | 0     | 0      |                    |                       |         |       |
| <                   | 4     | 4      |                    |                       |         |       |

<Totals>

|                     |       |        |                    |                       |         |       |
| BM II               | 84    | 33     |                    |                       |         |       |
| BM III              | 14    | 15     |                    |                       |         |       |
| PII/III Habitation  | 71    | 26     |                    |                       |         |       |
| PII/III Non-Habitation | 104 | 30 | |                       |         |       |
| Special             | 7     | 7      |                    |                       |         |       |
| Other               | < 11> | < 4>   |                    |                       |         |       |

291 115

* UG C-18-1 was inventoried as a P II/III habitation site, but when it sampled and collected it was found to be a BM III site. Another PII/I habitation site was then selected for collection.

<The Architectural Tree Ring Survey>

In addition to the surveys reported here, we surveyed, mapped and collected sites in selected canyons during 1974
Matson, Lipe, and Haase (Aug. 88) III-23 to obtain architectural and tree-ring information, as we mentioned previously (Matson and Lipe 1978; Lipe and Matson 1975). This work will be reported in detail later, although some of the dates obtained and preliminary results are used in this volume.

The sites studied were almost all in Grand Gulch, between Upper Grand Gulch and the mouth of Bullet Canyon. In addition to working in this section of Grand Gulch and the lower reaches of the canyons joining Grand Gulch, a portion of a small canyon on the east side of the mesa—McLoyds Canyon was surveyed. Moon House, a well preserved and often visited cliff dwelling, was the main focus of study in McLoyds Canyon.

<Field Procedures>

We spent the 1972 and 1973 seasons surveying and collecting quadrats in the five drainages following the design outlined above, as well as inventoring half of the canyon sites. As noted above, the quadrats were selected by simple random samples of each drainage, at a rate of 7 percent. This resulted in samples ranging from 9 to 22 quadrats per drainage and a total of 76 quadrats in all. The actual procedure used to select quadrats was to first to prepare a mosaic of aerial photos for the target drainage and to overlay a mylar sheet gridded with lines representing 400 meter quadrats. The drainage borders were then drawn on the overlay, the potential quadrats numbered and then random numbers drawn. The selected quadrats then had their
boundaries copied from the overlay onto the aerial photos, which had a scale of 1:5000. The boundaries drawn on the aerial photos were used to locate the quadrats in the field. The field procedure began by locating a corner of the quadrat by reference to the aerial photograph. Then, the sides of the quadrat were marked off on the ground, using compasses and measured ropes (and abundant flagging tape). The actual quadrat boundaries were then drawn on the aerial photo. The quadrats were surveyed by a crew of three to six, who walked abreast across the quadrat, from boundary to boundary, keeping their location by the use of compasses. The usual pattern was to have a crew of five and to have a spacing of 10 to 15 meters between crew members. Isolated artifacts were collected and their location noted, in estimated meters north and east of the quadrat’s southwest corner. Concentrations of artifacts or features were usually just noted and flagged during the search procedure. After the quadrat had been completely searched sites would be returned to, and mapping units laid out by measuring tapes and with the aid of compasses. These were usually 50 meter squares with each grid being mapped on a separate piece of graph paper. Major physiographic features were first noted and then each grid searched. An artifact found within such a unit was given a location number and placed in a plastic bag. The number was then recorded on the map showing the location. These locations are not points but might include a number of objects within one or two square meters. In areas of high artifact concentration, grids of 2
Matson, Lipe, and Haase (Aug. 88) III-25 by 2 or 3 by 3 meters were used instead. All features, such as hearths, slab-walled structures, or masonry rooms were assigned letters of the alphabet and were mapped separately at a larger scale. By these means all artifactual material in the quadrat was collected or mapped and the provenience noted, as for example, in Figure III-2.

In addition to artifactual information and features, other information was collected for each concentration or site. First-hand impressions about temporal position and cultural affinities as well as inferences about functions of structures were coded on forms (Figure III-3). Most of the environmental information was recorded on two forms, one of botanical variables (Figure III-4) and one of physiographic variables (Figure III-5). These two forms were completed for all sites and for all quadrats. After removing some rare and never-encountered plants from the forms, we had 72 usable botanical variables. These ranged from placement within a major plant community (pinyon-juniper, sagebrush, etc.) to identification of specific plants which were noted as present or absent for each site. Thus a wide variety of botanical information was collected for each site and quadrat.

The physiographic form included such information as type of surficial deposit, exposure, amount of dissection, and so forth. An important measurement was the depth of the caliche zone which limits the effective soil depth in many areas at Cedar Mesa. This depth was found by testing excavated soil with hydrochloric acid. The depth of soil is
Matson, Lipe, and Haase (Aug. 88) III-26
related to the sort of plants found; sagebrush, for
instance, was usually found in deep soil areas, although
some interesting exceptions occur.

In general all quadrats were completely mapped and
collected. Only 4 of the close to 300 sites found in the 76
quadrats were subsampled and in all four cases, dense
concentrations of trash were the contexts that were
subsampled. With these few exceptions, all artifactual
material found within each quadrat was completely collected.

The procedures involved in coping with sites that straddled
quadrat borders, however, were not consistent. For many of
our analyses only that portion of a site inside a quadrat
could be used (all analysis of variance approaches) but in
other situations, sites were the objects of comparison,
rather than quadrats. For these analyses, information was
need on all of a site, including the portions falling beyond
the quadrat boundaries. In some cases, the portion of the
site falling outside the boundary was collected, in other
cases it was not. In general, though, a feature was
completely collected if it straddled a quadrat boundary. In
all cases the material outside the quadrat was designated by
adding "A" to the site number, e.g., Bullet 2-3A, indicating
the outside-the-quadat portion of the third site in the
second quadrat in the Bullet drainage.

The drainage canyon inventory and collection procedures
were basically similar to the quadrat procedures but
involved a few differences which resulted from the different
setting. First, the searching procedures were different.
In a situation where one can walk back and forth across a relatively flat area, one can easily control the intensity of the survey; in a steep walled canyon, it is much more difficult. Generally the search was made by walking out ledges, down one side of a canyon and up the other. Such flat areas as existed on the canyon bottom were also searched. While the search intensity was more variable than in the quadrat search, the factors of abundant slickrock and protected shelters probably combined to make such sites as existed, more visible than on the mesa-top. Our impression is that more sites were probably missed in the canyon due to variable coverage than on the mesa-top, but that more sites on the mesa top were missed due to ground cover and natural erosion.

Because lithic source material occurred in seams of limestone which are exposed in the Cedar Mesa sandstone, various amounts of this material was found during the canyon inventory, both in the canyon washes, and along the sides of the canyon. On the mesa-top this material was collected, either as off-sites locations, or on sites, because it was brought to the mesa top by human action. During the canyon inventory abundant raw material was noted but not collected. In general, because of the focus on sites, and the problem of lithic source material, there is not a consistent category equivalent of the quadrat off-site material for the canyon survey.

A special form (Figure III-6) was developed for the canyon inventory. It was this form, with accompanying
Matson, Lipe, and Haase (Aug. 88) III-28 sketches that was filled out for all sites located, and was used as the basis for the preliminary classification of sites for sampling, as previously described. In addition, during inventory, botanical forms were filled out for every 400 meter stretch of the canyon, as was a water resource inventory form for every observed water source (Figure III-7). Since the location of water resources was thought to be important for the location of mesa top sites as well, this was important information for the quadrat survey and we will later refer to this information. In addition, L. Agenbroad (then of Chadron State College) carried out a survey of alluvial deposits in Upper Grand Gulch, Middle Grand Gulch and Bullet canyons, which complemented the canyon survey (Agenbroad 1975).

After the inventory was completed and the sites were classified, the canyon sample was taken and the selected sites returned to for mapping and collecting. The procedures involved here were essentially the same as those used on the mesa top. Differences involved the use of irregularly shaped mapping units, to conform to ledges on the canyon wall, and a variety of scale drawings, as well as plan views, where architectural remains were well preserved. The same location procedures, forms and so on were used as on the mesa top.

Throughout the canyon survey, we may appear to have veered from many of the basic principles of sampling and data comparablity that were at core of the overall project. While the quadrat survey inspected only a small fraction of
the overall ground surface of Cedar Mesa (less than 2 percent, and only 7 percent in the selected drainage) we ended up looking at a much higher proportion of the area's canyon sites, and up to 100 percent of the canyon sites in the selected drainages. In addition, up to 100 percent of certain site categories were collected. Overall, through the canyon survey of the five selected drainages, plus the architectural tree-ring survey, we looked at approximately 40 percent of the canyon sites in the study area.

Sampling theory works best when one is taking a large (in size, not rate) sample from a large population (sampling at a low rate). In the canyons we were forced to take small sized samples from small populations, which are poor conditions for sampling efficiency. Yet to compare and contrast the situation in the canyon bottom and mesa top, we needed further information about the canyons than was obtained during the quadrat survey. Under these conditions obtaining workable sample sizes becomes paramount, in spite of sampling "inefficiency". So in spite of our use of sampling theory, we recognize its limitations, at least in applications used in cookbook fashion in real life archaeological situations. In this case, in order to get sample sizes large enough for useful comparisons, we changed our sampling unit and sampled at a very high, "inefficient" rate.

<Analytical Procedures>

<Assumptions>
Once one has presented a rationale for making observations and described the field procedures, one is ready to compare the observations with the \textit{a priori} expectations. First though, assumptions about the nature of these observations need to be presented.

One key assumption has to do with uniformity over the years involved; that is, that the same processes operating today are those that occurred in the past. We do not mean that the environment or the sites have remained unchanged since the prehistoric inhabitants left; instead that the basic natural processes that determine the differences in the environment on Cedar Mesa and that are eroding archaeological sites today, are those that operated in the past.

We have placed emphasis on recording the present environment of geographical locations, both sites and quadrats. These attributes were not recorded in the belief that the environment was unchanged and that by recording them we were recording the exact prehistoric environment. Rather, we observed them in the belief that the environmental gradients that exist today on Cedar Mesa also existed in the past. For instance, if a certain site exists in a relatively dry location today, we assume that it was in a relatively dry location in the past, regardless of the climate present at that time. In other words the environmental gradients observed today are dependent on a number of factors that have changed in the recent past such as elevation, exposure, soil depth, and direction of
Matson, Lipe, and Haase (Aug. 88) III-31 prevailing wind flow. Overlaid on these basic geographic factors are others that have varied over the past two thousand years. These include cycles of erosion and aggradation and regional climate expressed locally as changes in length of growing season, and amount and seasonal distribution of precipitation.

Alluvial aggradation and arroyo cutting, which greatly change the nature of alluvial soil bodies, have alternated on Cedar Mesa, as elsewhere in the Southwest (Agenbroad 1975, Hack 1942). These processes undoubtedly resulted in differential attractiveness of the canyon environments for agricultural settlement at different periods. Because alluvial soils are only a small fraction of the arable soils on Cedar Mesa (Lipe and Matson 1975, Matson and Lipe 1978), alluvial cycles probably were not the primary factors controlling overall habitability of Cedar Mesa.

With regard to the upland aeolian soils we have argued previously that neither extensive degradation or aggradation on Cedar Mesa has occurred in the last two thousand years. During survey we did note local exceptions to this. Aggradation has taken place locally as have differing amounts of erosion. In general, however, deep soil areas today were deep soil in the recent past, and vice versa.

These arguments favoring "environmental relativism" beg the question of what differences did occur between present and past. Some effort was spent on this question (West 1978) and the results demonstrate that according to pollen analysis the same plants were on the mesa in the past as are
Matson, Lipe, and Haase (Aug. 88) III-32 there today. While our data do indicate a past environment similar to today in the past, they are not sensitive enough to sense subtle changes which may have been critical for agriculture. Thus we are unable to come up with any significant new information on this old question that is specific to Cedar Mesa. We will deal with this issue from the standpoint of evidence for regional climatic changes, several times later in this volume.

The other aspect of uniformitarianism we have assumed is in the structure of archaeological data. Previously we have mentioned that there is little evidence for either extensive erosion or aggradation with 1700 year old archaeological features often exposed at what must be within a few centimeters below their original surface of origin. We have also made reference to the general lack of ground cover. These two observations support the assumption that our surface collecting and mapping yield data representative of the contemporary archaeological record, and that this is in turn, representative of the original archaeological contexts.

Most Cedar Mesa sites are shallow, briefly occupied, single component sites. This means that the debris from the last occupation is not covering up earlier material of a different phase. Particularly for the mesa-top sites, if in the brief time span of occupation, site function changed it is not likely to have been buried by debris from later occupations. In short, given the nature of the environment, the stability of the soil, and the nature of the occupation
Matson, Lipe, and Haase (Aug. 88) III-33
our assumption of representativeness of the surface
collection appears to be well justified. Two major
exceptions are certain canyon sites and the mesa-top
Basketmaker III occupation, as will be described later.

Some independent confirmation of this assumption is
present in our limited excavations. In general what was
seen on the surface was confirmed in the excavation.
Deviations occurred in the nature of features found in
excavation as compared with those postulated from surface
evidence, but not in the general nature of the site. While
correspondence between surface and subsurface material has
been presented before (Redman and Watson 1970) generally the
existence of correspondence has been demonstrated in far
inferior conditions than found on Cedar Mesa.

One can readily see the correspondence between surface
and subsurface when masonry rubble and well preserved
sandstone slab features are present on the surface. But for
many sites, such features are either not clear or are
lacking, forcing reliance on surface artifactual material
for functional inferences. Site location and site dating
information, in most cases, are not subject to
transformations that would make them unclear. However,
accidents of erosion, location of tree duff and prior
occurrence of casual collecting all make surface artifact
information less "representative" of original site
inventory, than is excavated of subsurface material. So
there is a hierarchy in representativeness here, with well
preserved structures at one end, and surface small artifacts
Matson, Lipe, and Haase (Aug. 88) III-34 at the other. Unfortunately, it is the latter that are most readily and regularly observed on the typical Cedar Mesa site.

For most of our work, we have not assumed that the spatial organization of surface material represents the original organization of the archaeological context. Dohm(1981) gives a detailed discussion of this question and an argument that, in general, organization is preserved, but this is not essential for most of our work. Further, in our discussion of R mode analysis, (later in this chapter) we discuss how this assumption can be avoided in a critical case.

<Analysis of Variance>

Much of the following analysis is carried out through the use of analysis of variance, comparing two or more populations. The parametric tests usually associated with this analysis is the t-test and f-test family at the simplest level, ranging up to very complicated multi-way multivariate procedures. We have chosen to use the Wilcoxon[-Kruskal][-Wallis series of non-parametric analogues for reasons that will become apparent. But first the nature of what we have sampled and what the sample units are, needs to be clarified.

The basic unit that was sampled is the quadrat and it is groups of these that are compared in most tests. Not only is there no practical way of generating lists of sites for the mesa as a whole, but even if one had that information, one would still not have information on density of sites and
Matson, Lipe, and Haase (Aug. 88) III-35 where they are <not> located. So even though our interest is in archaeological information, which would lead one to expect sites and off-site artifacts to be the units of analysis, the units in most cases of analysis of variance tests were quadrats, with the archaeological information used as quadrat <attributes>. For many other non-analysis of variance purposes, however, sites are used instead.

In the comparison between drainage canyons and mesa tops we are comparing a random sample of sites (drainage canyons) and one of quadrats. This is very difficult to do and instead, the <sites> found in the quadrats were compared with those in the canyon. Since, strictly speaking, the mesa-top sample is one of quadrats, this diverges from the procedure described above. However, the sample of mesa top sites is a representative and random sample since the quadrats were random. One problem is that sites (and artifacts) occur as clusters within quadrats, so the <size> of the sample is the <n> of the quadrats, not the <n> of the sites recorded on the mesa-top.

This problem has been referred to as cluster sampling (Thomas 1975, Mueller 1975). Thomas (1975:79) suggests a way around this for abundant items by the use of cluster sampling formulas and ratio-means. Note, however, that even his own examples, using his most abundant variables, violates the rules he presents. Other guides suggest more stringent rules (Cochran 1963) which mean that except when one is dealing with very abundant variables such as total debitage or total ceramics, this approach is not likely to
Matson, Lipe, and Haase (Aug. 88) III-36
be appropriate when using quadrat samples. We may conclude
that cluster sampling, while conceptually proper, is not
something that can usually be applied in this fashion for
quadrat sampling, as the usual numbers are of the wrong
order, being far too small. So, for most cases, using
quadrats as the elements is not only reasonable but the only
practical way to proceed.

Fortunately, the theoretical problems of comparing
mesa-top and canyon samples on Cedar Mesa, are usually not
important in practice because the \( n \) of sites is usually
below the \( n \) of quadrats when one is dealing with any one
type of site. Further, the grab sample effect (sites within
a quadrat are liable to be more similar to each other than
sites in the population as a whole) is not a major concern
because it is unusual for more than one site of any
particular kind used in these comparisons to be found in any
single quadrat. Thus the quadrat same of the specific site
types used in comparisons between the mesa-top and canyon
approximates a random sample of that site type with an \( n \)
of the number of sites.

To summarize, the drainage canyon sample is of sites as
elements, while the mesa-top is of quadrats as elements.
The way we compared the two is by reducing both to the same
elements, i.e., sites. But the quadrat sample is really a
cluster sample of sites. However, the numbers for site
comparisons can not be corrected by cluster sampling
formulas. On the other hand the \( n \) of sites for the mesa
top is usually less than the \( n \) of quadrats in these
Matson, Lipe, and Haase (Aug. 88) III-37 comparisons, so the inflated <n> does not occur in the first place. These points will be examined carefully when the few actual comparisons we made are described.

When we examine the distribution of the quadrat variables or attributes of interest, whether sites or artifacts we find, they are typically skewed to the right—that is, low numbers are more common than high numbers, producing a one-tailed curve. This means the variables are not normally distributed. This situation is not uncommon in archaeology (Matson 1976:104) and is the usual situation in quadrat samples (Matson 1971, True and Matson 1974, Thomas 1975, Matson and Lipe 1975). This fact is often ignored in archaeology, as if it will go away if it's not mentioned.

While the distributions of most archaeological variables per unit area in this project do not approximate the normal distribution, much of sampling theory is based on the assumption that the distribution of repeated <estimates> of the population values does approach the normal distribution (Cochran 1963:11-16). This assumption is easily verified for large samples of normally distributed variables. However, the utility of this approach becomes less as the population values depart more from the normal curve and sample size becomes smaller. Some sampling texts do discuss this problem in some length (cf. Cochran 1963: 38-44) but nevertheless continue to use the assumption of normality. It may well be that such assumptions are well employed in sociological surveys where the samples are large and
Matson, Lipe, and Haase (Aug. 88) III-38 population variables may be approximately normally distributed. The utility of this assumption for sociological sampling has, however, been questioned (Kempthorne 1969:674-5). In fact, referring to the accuracy of estimates of standard error, an often used crucial parameter, an introductory text states:

"That is, however, a more difficult problem, about which we shall say only that, under realistic assumptions about distribution of the measurements X and the population value V, a rather large sample is required to produce satisfactorily precise estimates"(Hodges and Lehman 1964:250).

In short, the assumption of normal distribution of population estimates is likely to be incorrect for archaeological survey sampling where the sample size is moderate and the distributions of population variables are non-normal(as is certainly the case with the Cedar Mesa Project). Thus two questions arise: what are the effects of improperly using such an assumption and what possible alternatives exist?

Three common uses of the standard error are comparing the precision of sampling methods, estimating the size of sample needed and estimating the precision attained in a sample by establishing confidence limits (Cochran 1963:23). It is quite clear that in our case assuming the "...use of the 'normal' formula for appraising the variance of S#2 (standard error) may give a very misleading impression of the stability of S#2 "(Cochran 1963:44). For Cedar Mesa in
Matson, Lipe, and Haase (Aug. 88) III-39
the three uses above, the precision of sampling methods
would be over estimated; the estimated sample sizes would be
too small and the confidence limits would be too narrow.
To see what the effects would be using Cedar Mesa data, a
series of samples were drawn from simulated data whose
population parameters were similar to those of the Cedar
Mesa samples (SAMSIM, Matson 1975a,b). The mean standard
error was calculated and the Student's t distribution was
used to calculate 95 percent confidence intervals. (That
is, intervals in which the population mean would be included
approximately 95 times out of a hundred if the assumption of
normality was correct.) While the Student's t distribution
is not exactly the same as the normal distribution, it is
closely related and recommended if the sample size is less
than 60 (Cochran 1963:50). In the run which showed the most
development from the expected on the basis of normal
assumptions, in 7 out of the 20 samples the population mean
lay outside of the calculated interval, instead of the
expected one in twenty. In another, the mean lay outside of
the interval 17 times out of a hundred. Even when the
number of means inside the interval was about 95 out of 100,
the intervals were often onesided. For example, in one case
the mean lay outside of the interval only \( \frac{x}{4} \) times out of 100,
but it was on the lower side in all seven cases (for
examples see Table III-2). Thus, with simulated data
similar to those from on Cedar Mesa, the use of "normal" (in
this case Student's t) approximation would lead to incorrect
inferences in many cases. Presumably this would be the case
Matson, Lipe, and Haase (Aug. 88) III-40 with other regional sampling projects; at least with any we are aware of. Errors would thus occur not only in establishing confidence intervals, but by extension, sample sizes and comparative precision of sampling methods.

There is, however, an alternative to assuming normal distribution of population estimates that works well with non-normal as well as with normal distributed data, that of the use of Binomial Confidence limits (Cochran 1963:58-59, Conover 1971a: 110). Here assumptions are not made about the shape of the distribution of the estimates but the probability of obtaining a set of "successes" or "failures" with only the ranks of the data being used in the calculations. Then the values of the data are used to replace the ranks to actually delimit the confidence interval (see Table III-3 for examples).

When the simulations were repeated using SAMSIM and binomial confidence intervals the results (see Table III-2) show that the binomial confidence intervals, are conservative and show little or no tendency for skewness in errors, in contrast to the intervals based on the standard error. Many other simulations were run in addition to those in Table III-2, including ones with data distributed in near normal shape, and the binomial confidence intervals repeated this pattern consistently. In general, then, one can use this approach to calculate valid confidence intervals on population parameters of non-normal as well a "normal" data. This alternative to the widely used normal theory is only a partial alternative, however useful it may be. We now
Matson, Lipe, and Haase (Aug. 88) outline some of the drawbacks involved in this approach and solutions to certain of these problems. Inspecting Table III-3 which shows the summary site statistics for the Bullet drainage quadrats, one can see some of the potential problems where the median appears to be unstable. We have actually found the median to be generally a very stable parameter, provided that the distribution is relatively continuous in the region of the median, or if it occurs away from any edge of a change, say, from 1 to 2, in a step distribution. This latter case can be visualized on Table III-3 for site medians, if the median is 2 but close to the step to 1. The Basketmaker II site medians might be a case in point. With repeated sampling it would sometimes be 1 and sometimes be 2, a difference of 100 percent. In the same situation the mean would vary less, e.g., 1.5 to 1.8, a much smaller percentage difference. While this kind of problem is less likely when looking at artifact counts, which are generally more numerous, the "step" problem may occur when working with infrequent items having abundant ties. It can sometimes, but not always, be avoided by going to another quantile, such as the third quartile. (Binomial confidence intervals are easily calculated for any quantile, that is, any percentage class). Even in these situations of unstable quantiles, the binomial confidence intervals appear to hold up.

TABLE III-2
Comparison of "Normal" Confidence Intervals on Mean with "Binomial" Intervals on Median.
Simulated data, sample size is 10, 100 samples.

<table>
<thead>
<tr>
<th>95% &quot;Student&quot; Intervals</th>
<th>98% &quot;Binomial&quot; Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means within</td>
<td>Medians within</td>
</tr>
<tr>
<td>85</td>
<td>97</td>
</tr>
<tr>
<td>Means outside</td>
<td>Medians outside</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Means low side</td>
<td>Medians tied low side</td>
</tr>
<tr>
<td>15</td>
<td>2*</td>
</tr>
<tr>
<td>Means high side</td>
<td>Medians tied high side</td>
</tr>
<tr>
<td>0</td>
<td>1*</td>
</tr>
</tbody>
</table>

Simulated "Cedar Mesa" data, 400 meter quadrats (see Table III-3) sample size is 21, 100 samples.

<table>
<thead>
<tr>
<th>Means within</th>
<th>Medians within</th>
</tr>
</thead>
<tbody>
<tr>
<td>83</td>
<td>99</td>
</tr>
<tr>
<td>Means outside</td>
<td>Medians outside</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>Means high side</td>
<td>Medians high side</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>Means low side</td>
<td>Medians low side</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* These are medians on the boundary of the confidence intervals.

**TABLE III-3**

Bullet Drainage Summary

Site Statistics for Quadrats(n=21).

<table>
<thead>
<tr>
<th>BM II</th>
<th>BM III</th>
<th>P II-III</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

Site Medians
Matson, Lipe, and Haase (Aug. 88) III-43

Site Interquartiles 0-2 0-2 1-4 0-2 4-9

Median 90% Binomial

Confidence intervals 1-2 0-1 2-4 0-2 4-8

Site Means 1.67 .95 2.90 1.43 6.81

Total artifact

Medians 147 50 525 4 1230

Total artifact

Interquartiles 0 - 508 0 -312 382-1503 0-17 721-2110

Median 90% Binomial

Confidence Intervals 41-228 0-278 390-1132 0-10 956-1674

Artifact Means 340.8 235.2 929.3 9.0 1514.3

While the median is in many ways a very useful measure of central tendency, it is not altogether a replacement for the mean. In a symmetrical distribution the mean is both an estimate of the median and long run <expected> value because it incorporates the total value of the population (or sample). While in the one-tailed distributions the mean has little value as a central tendency, it is a measure of normed total or expected value. Unlike the mean, the median is not sensitive to the size of the extreme values. This is generally desirable for extremes are usually rare and thus poor subjects for sampling. In most cases, ignoring the extremes is reasonable but in some cases much of the
Matson, Lipe, and Haase (Aug. 88) III-44
archaeological material is in the extremes. In this
situation, there is sufficient reason for using the mean as
a population parameter.

There are then basically three reasons for using the
mean: because of most readers' familiarity, because the
median is unstable in some step distributions and because
the mean gives a normed total or expected value. While we
have pointed out the unsuitability of using normal theory
for sampling on Cedar Mesa, it is possible to calculate
confidence intervals for means without assuming a normal
distribution. Unfortunately some of these involve assuming
some other population distribution such as the Poisson
(Lindsey 1970), and inspection of our data suggests that
they do not fit into any common distribution, at least
untransformed. Since we did our analysis, however, the
empirical "Boot Strap" approach (Efron 1981) has also been
developed which can be used to some degree. Another
"drawback" to using non-parametric confidence intervals is
that the probabilities do not usually fall on the
traditional 0.001, 0.01 and 0.05, forcing one to use
irregular intervals that are close to the traditional ones.
We do not see this as more than a minor problem. The
binomial confidence intervals are also wider than those
calculated by normal distribution theory, which is to be
expected, since the additional assumption of symmetry is not
used (and this assumption turns out to be invalid with this
kind of data). Not assuming symmetry is not a drawback when
the distributions are not symmetrical. There are also ways
Matson, Lipe, and Haase (Aug. 88) III-45
other than using binomial theory of calculating quartile
intervals which are said to usually give narrower intervals
(Conover 1971:216)

The above problems are not very important compared with
the alternative which is to use statistics of uncertain
value because their assumptions are widely violated.
Consequently we chose to use non-parametric confidence
intervals in this work. Non-parametric sampling and
statistical tests are rapidly expanding fields, particularly
as used in medicine, and solutions to some of our dilemmas
undoubtedly already exist and, it is hoped, will become more
generally known shortly.

Since the classic parametric analysis of variance tests
are clearly inappropriate with these data sets because they
assume near normal distributions, we used the
Wilcoxon-Kruskal-Wallis family of nonparametric tests
(Bradley 1968, Conover 1971a). These tests are based on the
ranks of the sampled values and so are distribution free.
It might be noted that the asymptotic relative efficiency
(A.R.E.) of Wilcoxon-Kruskal-Wallis family, with respect to
the familiar t test–f test family against the location
alternative, is at a minimum .864 and at a maximum is
arbitrarily high (Bradley 1968). These figures indicate
that at a maximum only a 13.6% larger sample is needed for
the Wilcoxon family to reject the null hypothesis at the
same level as the t-f tests. Conversely there are
situations where no matter how much larger the sample is
made for the t-f test family, they will not reject the null
Matson, Lipe, and Haase (Aug. 88) III-46
hypothesis at the same level as the Wilcoxon-Kruskal-Wallis
tests. Experiments with small samples indicate that these
figures are good approximations for small samples.
Incidentally, even in the normal theory case, that is, where
the data do follow the normal distribution, the A.R.E. of the
Wilcoxon test is still 94.5%. Thus one has little to lose
and much to gain by using this family of tests for testing
statistical significance.

While we have discussed some of the disadvantages in
having non-normal data and using quantiles instead of the
classic mean and variance, equivalent drawbacks do not
appear to exist in using this family of tests, even in
situations where the data have a normal distribution. In
terms of intuitive understanding, power, and ease of use,
these tests compare very favorably. It is true that the t
and f test have been generalized into more complicated
multiway, multivariate extensions while the non-parametric
equivalents are either cumbersome or, in practice,
unavailable. We have not needed to use these more complex
tests and judge both families to have assumptions that
cannot usually be met with archaeological data. In short,
there do not seem to be significant disadvantages for using
the Wilcoxon family of tests. In most important aspects,
these tests are to be preferred even if one has a choice,
which was clearly not available to us.

<Evaluation of Sampling Design>

The choice of 400 meter square quadrats turned out to be
very successful. In forested regions Matson noticed significantly fewer problems than with the 500 meter quadrats he had used previously (Matson 1971, 1974). Of the 76 quadrats surveyed on Cedar Mesa only 6 did not have "sites" as we defined them, and only one was totally barren of all archaeological material. The 400 meter quadrat size has since been used elsewhere with good success (Pokotylo 1978, Matson, Ham and Bunyan 1984, Matson, Magne, Ludowicz and Pokotylo 1980).

The sampling design, as developed, is a compromise and once we obtained our sample we were interested to see what variables were well sampled and what alternative sizes of quadrats and samples would have produced. As might be expected, common variables were oversampled and rare ones, such as special kinds of sites, were sampled at minimal levels. As analysis progressed and more and more detailed questions were proposed, the sample showed more weaknesses. In order to evaluate the sample more rigorously, Matson developed the previously mentioned simple sampling simulation program (Matson 1975a,b) to run empirical tests of Cedar Mesa and other sampling schemes (Matson and Lipe 1975).

From the previous discussion, it should be clear that most lithic and ceramic types could be adequately sampled with much smaller quadrats or at a lower sampling rate. On the other end, composite variables such as specific "site" types may only be minimally represented. To explore the reliability of sampling of such composite variables we
Matson, Lipe, and Haase (Aug. 88) III-48 conducted some sampling experiments based on the samples that had been collected during the first field season.

From the beginning we considered that very "rare" items such as "defensive" sites would be "sampled" in some other fashion than the above. We hoped, however, that such items as "unit Pueblos" (ordinarily, a kiva and several associated room structures), while rare relative to other site varieties, would be common enough to be sampled reliably. As one might expect, after sampling only two drainages, we were unsure of the reliability of our sampling with respect to the "unit Pueblos." This problem appeared to be more acute in the Bullet drainage where the apparent density of this site type was much lower than in Upper Grand Gulch, the other drainage sampled in 1972.

In order to obtain an empirical check on the reliability of these estimates we decided to assume that our sample was representative and to generate a target population with the same population parameters. This artificial population was assumed to be an approximation of site distribution for a complete drainage. Because of low occurrence we chose the mean as the parameter of interest for the abundance of unit type pueblos. We then repeatedly sampled this artificial population at the rate that had been used in the field. Thus a 7 percent sample of 400 meter quadrats was repeatedly simulated. (This approach resembles the more sophisticated and rigorous "Bootstrap" approach which has been developed, in the main, since the Cedar Mesa Project began.) In our preliminary analysis only two unit pueblos were found in the
actual Bullet drainage sample. In a simulation of 20 such samples, only one sampling of the area did not have any unit pueblo and only one sample had more than three units (see Table III-4). Even though the item was very rare, found in only 10 percent of the quadrats sampled, most of the simulated samples were in good agreement with one another and with the assumed "actual" distribution of this type of site. Thus our faith in our sample of unit Pueblos for the Bullet drainage was considerably strengthened.

Of course the real distribution of sites may be different and our sample may be "unrepresentative." Still our best estimate of the population parameters are our sample parameters and if sampled parameters from the assumed population were unstable we would have a different view of the reliability of our sample.

In the Upper Grand Gulch drainage, the unit type pueblos were more frequent, although fewer quadrats were collected as the drainage was smaller than Bullet. Again, the question of reliability was checked by further sampling simulation, with the field-sampled distribution assumed for the drainage as a whole. Again the reliability was better than expected (see Table III-5) with only 6 samples out of 100 lacking any unit pueblos and only 5 samples having three or more units than expected.

TABLE III-4

Comparison of Simulation of 400 Meter Quadrat Samples of "Unit" Type Sites from Bullet With 800 Meter Quadrat
Matson, Lipe, and Haase (Aug. 88) III-50
Sampling Simulation

(20 samples, each of 21 quadrats)

\[
\begin{array}{cc}
800 \text{ Meter} & 400 \text{ Meter} \\
\hline
\text{Sample mean (of means)} & .122 & .129 \\
\text{Sample median (of means)} & .12 & .14 \\
\text{Interquartile (6,15) (of means)} & .95,.143 & .10,.14 \\
\text{Low (of means)} & .067 & 0.0 \\
\text{High (of means)} & .17 & .29 \\
\end{array}
\]

Note: Population mean is .12 per 160,000 sq. meters.

---

TABLE III-5

Upper Grand Gulch 400 Meter Quadrat
"Unit" Type Site Simulation

(100 samples, each of 9 quadrats)

\[
\begin{array}{c}
\text{Sample mean (of means)} \quad .445 \\
\text{Sample median (of means)} \quad .375 \quad ( .375, .500 ) \\
\text{Interquartile (25, 76)} \quad .25, .75 \\
\text{Low (of means)} \quad 0.0(6) \\
\text{High (of means)} \quad 1.0(1) \\
\end{array}
\]

Note: Population mean is .46. ( ) enclose 95% confidence intervals of median. { } indicate number of samples with that value.
Thus even for site types which are much less common than most of the variables we are interested in, the techniques employed (sampling rate, quadrat size, and collection procedures) give us reasonably precise estimates of their overall occurrence, although not of their geographic locations within the individual drainages. Clearly, if we were only interested in the number of unit Pueblos, the present sampling design and collecting procedures would be woefully inefficient. A much less intensive search of, say 800 meter quadrats with collection only of unit type sites would be much more efficient, as these large quadrats could be searched and collected in the same or less time, than are the 400 meter quadrats are under the present scheme.

We simulated a comparison of this procedure using the same sample size as used in the project (a sampling rate of 28 percent as compared to 7 percent because the 800 meter quadrats have four times the area). We again based our population on the survey results of the first season and expanded it to fit the new conditions (Table III-4). Because the median in all cases would be 0 we are again using the mean as the population parameter of interest. With a run of 20 samples the means of the 800 meter quadrats ranged from .07 to .16 (mean of population for unit pueblos per 160,000 sq. meters (the area of a 400 meter quadrat which we used as our "standard") is .12). This is compared to a spread of 0.0 to .29 (mean of the population is .12) for the 400 meter quadrat simulation. Thus, as expected, the 800 meter quadrat scheme would be much more precise for
Matson, Lipe, and Haase (Aug. 88) III-52
the same or less money in terms of the distribution of unit
pueblos in the Bullet drainage.

If some variables of interest were undersampled and
could have been sampled more efficiently using some other
scheme, the opposite is also true. Some common variables
were oversampled and could have been more efficiently
sampled by using a larger number of smaller quadrats but
with a smaller amount of area being surveyed. Such a
variable is the amount of Pueblo II-III artifactual
material. Again, assuming the sampled artifact density was
true for the population we simulated 100 samples of 400
meter quadrats at a 7 percent rate (Table III-6). We then
simulated 100 samples of 200 meter quadrats at a 2.4% rate
using the same assumptions about the population. Note that
the population parameters are slightly different in the two
populations. In "expanding" the 400 meter quadrat
population to fit the larger population of 200 meter
quadrats, in an attempt to add realism we added and
subtracted random values and ended up with slightly higher
population values for the 200 meter quadrat case. This step
was considered to be much preferable to generating the new
population by taking the 400 meter population, dividing each
sample element by 4 and repeating this population 4 times
which would result in an overly homogenous population since
it would not take into account the clustered nature of the
data (See Matson and True 1974).

As seen in Table III-6, the 200 meter quadrat samples are
substantially more precise than the 400 meter quadrat
samples with a definitely smaller interquartile range (1039 to 1362 (a difference of 323) compared to 969 to 1365 (396)). This increase in precision is due to increased sample size (30 as compared to 21) in spite of the decrease in the area sampled.

Table III-6

Comparison of Simulation of 7% 400 Meter Quadrat Samples of Pueblo II-II Material with 2.4% 200 Meter Quadrat Samples (100 Samples each)

<table>
<thead>
<tr>
<th></th>
<th>400 Meter</th>
<th>200 Meter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number in sample</td>
<td>21</td>
<td>30</td>
</tr>
<tr>
<td>Pop. Mean (per 160,000 sq. m.)</td>
<td>1202.2</td>
<td>1231.3</td>
</tr>
<tr>
<td>Population median</td>
<td>970</td>
<td>1010.1</td>
</tr>
<tr>
<td>Sample mean (of means)</td>
<td>1172.7</td>
<td>1216.1</td>
</tr>
<tr>
<td>Sample median (of means)</td>
<td>1130</td>
<td>1130</td>
</tr>
<tr>
<td>Sample interquartile (of means)</td>
<td>969, 1365</td>
<td>1039, 1362</td>
</tr>
<tr>
<td>Sample low (of means)</td>
<td>676</td>
<td>748</td>
</tr>
<tr>
<td>Sample high (of means)</td>
<td>1957</td>
<td>1967</td>
</tr>
</tbody>
</table>

A comment is in order here. This example is a case in point of the effect of decreasing sample unit size and thus increasing sample <n> while not increasing sampling rate (Plog 196, Matson and True 1974), which is denied by Hole (1981). The gain in precision is significant. Here we find
Matson, Lipe, and Haase (Aug. 88) III-54

a sampling rate of 2.4 percent is more precise than one of 7 percent because of this effect. Similar results can be obtained in other cases as well.

Many other simulations were run, but the results were basically repetitive of the points made above. In summary:

For many of the variables under consideration much more efficient sampling schemes could be developed. Since, however, we are concentrating on a wide range of variables, including very common and small scale ones as well as rare and large ones, we needed a sampling design that offered general usefulness, rather than a high efficiency for one or few variables. Given our interests, a generalized sampling design would be the most efficient, offering relatively reliable data on a number of variables. The alternative solution, to design a number of different sampling techniques for specific variables, would seem to be unduly complicated and probably unworkable.

For variables that have very limited distribution, however, such specialized sampling schemes might well be used in conjunction with the above. For instance, defensive sites, which are unlikely to be recovered by our sampling procedures, could be sampled in other fashions. Another case is canyon sites, for which we did develop a very different sampling scheme.

In general, our sampling design held up well under these tests, and these checks indicate the sample is reliable. At the same time because of the nature of survey data, the form of the observations are often cumbersome and not those which
Matson, Lipe, and Haase (Aug. 88) III-55 would be most useful, i.e., the use of quadrats as the sampling elements rather than sites. The proof of any pudding is in the eating; the true evaluation of the sampling design will be in the later substantive sections.

<Data Reduction>

Analysis of variance type statistics presuppose that the units being can be placed in homogenous classes for some attributes, while varying in others. An example will make this clear: we have hypothesized that Basketmaker II habitation sites are more abundant on the edges of the mesa than elsewhere. Analysis of variance requires dividing up the quadrats according to canyon rim or non-rim locations and contrasting the abundance of Basketmaker II habitation sites in the two subpopulations of quadrats. One aspect of our use of data reduction is to develop classifications, e.g., to decide which sites are Basketmaker II habitation sites necessary to make this sort of analysis.

The use of data reduction techniques to develop such groups agrees well with other aspects of our overall approach. First, we conceive of archaeological entities as generally polythetic, rather than monothetic. That is, entities such as artifact types, tool complexes, types of features, etc., are considered to be groups or populations of more or less similar entities; the members of such groups share a large number of characteristics, but no single characteristic is both sufficient or necessary for inclusion of a member in a group (Beckner 1958; Clarke 1968; Sokal and
Matson, Lipe, and Haase (Aug. 88) III-56
Sneath 1963). If one wants to characterize a group by an
abstracted list of characters, these must be expressed in
terms of frequencies in which they appear in a group, or in
terms of the tendencies for characteristics to associate
with one another, as measured by their co-occurrence or lack
of occurrence with group members, the members being taken
one at a time. The use of numerous variables in defining
polythetic entities does not result in untoward splitting of
groups, as in hierarchical monothetic classification, with
the result that polythetic entities may have quite high
information content. Use of a large number of variables, as
we propose in some cases, does generally require that
explicit quantitative methods be used, and this in turn
generally requires the assistance of a computer.

Second, we have collected a large amount of information
about each site. If we were to use an <a priori>, deductive
classification scheme we would not be making use of this
material. The only way such kinds of information can be
handled wisely is to use it to develop classifications with
"high information" content. This generally requires formal
data reduction techniques.

Third, the use of such techniques results in an
objective classification, that can be reproduced. While
there are many alternative objective classifications, any
objective classification has the advantage over either
intuitive or <a priori> deductive classifications of being
objective, or having a necessary relationship with the data.
The pigeon holes created may well be suboptimal but the
Matson, Lipe, and Haase (Aug. 88) III-57
data will be not forced into inappropriate holes. Further,
by applying slightly different methods, we can evaluate the
reliability of classifications, and whether other
alternatives are as suitable.

Finally, our approach is one of combining deductive and
inductive procedures. By using an inductive pattern seeking
approach to classification, we ensure that there is a
connection between our data and the ideas we are testing.
Data reduction techniques thus accord well with our general
approach.

While one use of data reduction techniques is to develop
classifications for use in hypothesis testing, another and,
equally important, use is for pattern searching, to
understand systemic structure. How many kinds of sites are
there? What are the kinds of sites? What are the major
axes of differentiation among sites? And how are these
aspects related to location? These kinds of questions are
all pattern seeking questions and are well answered by data
reduction procedures.

The techniques we have chosen to use are clustering and
scaling. We have argued for the use of these in archaeology
for some time (True and Matson 1970, Matson 1974, Lipe and
Matson 1971, Matson and True 1974) and they have been used
by many others as well (Cowgill 1968, Johnson 1968, Hodson
and Doran 1975). In some ways the development of clustering
and scaling parallels that of non-parametric statistics; the
original claim was not for superior results but that the
classical methods, such as factor and discriminant analyses,
had assumptions that severely limited their use. The early clustering and non-metrical scaling methods were made possible by the computer. As time has gone on the differences between these newer techniques and the classical multivariate data reduction techniques have become less distinct. This is particularly so with multidimensional scaling, the version we will use the most is closely related to principal components analysis. Now the claim is often made (Ref??) that these methods are superior to the classic methods, even when the assumptions of the latter are met.

We will use only a few of the very large number of clustering and scaling techniques that are available. In many cases different algorithms produce the same result. In other cases, there are no clear reasons for preferring one or the other. The set that we work with and the way we use them have the advantages of generality and a unified conceptual approach.

The first step we use in either clustering or scaling is to calculate the initial similarity matrices using <distance> measures. The basic similarity matrix measures the amount of likeness between entities in a pairwise fashion. By using distance measures we can visualize this similarity in terms of space, with identity being the same spot in space and great differences being large distances. The similarity matrix then represents a <configuration> of locations in space with similarity being indicated by closeness in space and difference by those entities located further apart.
Matson, Lipe, and Haase (Aug. 88) III-59
A simple example will demonstrate what we are describing. Assume four sites, A, B, C, and D, with four types of pottery present or absent: plain gray, corrugated gray, black-on-white and red-on-orange. The initial data matrix might look like Table III-7.

Table III-7
Hypothetical Data Array: A, B, C, and D are sites; + and - indicate presence or absence.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Plain Gray</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>Corrugated Gray</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>Black-on-White</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>Red-on-Orange</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

From this we could calculate a similarity matrix using a distance measure, called Jaccard’s distance or Jaccard’s complement. Given two entities, I and II, this measure is defined as:

\[
I \left( \frac{b \cdot c}{a + b + c} \right)
\]

where

\[
egin{array}{cc}
+ & - \\
\hline
+ & a & b \\
- & c & d
\end{array}
\]


For example comparing sites A and B, a = 1, b = 1, c = 0, d
Matson, Lipe, and Haase (Aug. 88) III-60
= 2 so that the distance is .50.

Note that the d cell, that of absences common to both
sites, is not used in this coefficient. We have found this
feature to be an advantage, particularly when working with
sites of different sizes. Jaccard's distance then counts
positive matches as points of similarity, mismatches as
points of dissimilarity, and ignores negative matches or
common absences (by not including d in the equation).
Another important aspect of Jaccard's distance is that it is
metric. We will point out the significance of this when we
discuss scaling.

The complete distance matrix is as seen in Table III-8.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.50</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1.0</td>
<td>0.75</td>
<td>0.0</td>
</tr>
<tr>
<td>D</td>
<td>1.0</td>
<td>1.0</td>
<td>0.33</td>
</tr>
</tbody>
</table>

This resulting distance matrix then can be clustered or
scaled. Clustering attempts to cluster entities into
classes, in this case asking the question of what sites
belong together, while scaling searches for the
dimensionality of the configuration, looking for the basic
factors determining the differences between sites.
Matson, Lipe, and Haase (Aug. 88) III-61

There are many different clustering algorithms, but we prefer those Sneath and Sokal have designated as S.A.H.N. (1974). Within this group, the three we use most often are Farthest Neighbor or Complete Linkage, Average Linkage (unweighted pair group method) and Ward’s error sum of squares. For many years, average linkage of some sort was the technique preferred (Sokal and Sneath 1963) but recently Wards and Farthest Neighbor have become more widely used.

All these hierarchical clustering techniques can be thought of as proceeding in the same way, first searching the matrix for the closest entities, linking them together and then updating the matrix and repeating the search stage.

It is only in updating the matrix that they differ.

In general we prefer the use of Farthest Neighbor or Complete Linkage. It is conceptually simple and theoretically sound, and does not have the chaining problems of Nearest Neighbor or Average methods. It is also indifferent to non-monotonic changes in the similarity matrix. This means that measures that give the same rank order of distances, such as squared and unsquared distances, or closely related coefficients such as Dice and Jaccard’s, give the same results. This is not true of Ward’s. Ward’s method has been evaluated as superior by some, but our impression is that these tests have been biased. A preference for Ward’s may be due to its reliance on squared error sums which make the clusters appear to be more discrete than do other methods. In expressing our preference for Farthest Neighbor, we do not mean that it is
Matson, Lipe, and Haase (Aug. 88) III-62
without drawbacks. There is one situation where it gives poor results, which we will discuss in the example below.
Further, we systematically compare Farthest Neighbor results with the other two techniques, Average and Wards, in our analyses.

In our example the closest entities are sites C and D, joining together to form a cluster at a distance of 0.33. How shall C,D be grouped with the remaining sites? In Farthest Neighbor clustering the similarity between any two clusters is represented by the farthest distance between any members in the two clusters and the matrix is updated accordingly. The distances between Site A and Site C and D are both 1.0 so that this case the similarity between our new cluster of C,D and site A is 1.0. The distances between site B and Sites C and D are 0.75 and 1.0. Using the Farthest Neighbor criterion, the longer distance is used, indicating a similarity of 1.0 between Site B and cluster C,D. The updated matrix is listed in Table III-9.

Table III-9

Recalculated Distance Matrix

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C,D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.50</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>C,D</td>
<td>1.0</td>
<td>1.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
The lowest distance value in this matrix is 0.50, the
distance between sites A and B so the next clustering step
is joining A and B at 0.50. Since the remaining values in
the matrix are 1.0, which is the maximum value that
Jaccard's distance takes, it is clear that the distance
between clusters A,B and C,D is 1.0, resulting in the
dendrogram seen in Figure III-9. This dendrogram would be
interpreted as showing two clusters of sites, A,B and C,D.

Table III-10

Hypothetical Data Array: A,B,C,D,E and
F are sites; + and - indicate presence or absence.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Plain Gray</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. Corrugated Gray</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3. Black on White</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4. Red on Orange</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

What would happen if there was another group of sites E
and F (Table III-10) joined to the rest at 1.0 as in Figure
III-9? How would one interpret the relationship between the
clusters? Which clusters are closer to one another and
which are further apart? With Farthest Neighbor clustering
this could not be done. Where separate clusters have
members differing by the limit(1.0), as in this case, the
Matson, Lipe, and Haase (Aug. 88) III-64 relationship between the clusters, is also at the limit, even if other members of the two clusters vary in their distances from one another. Thus when one is interested in higher order relationship between clusters, and one has this kind of structure in the data, the Farthest Neighbor method is not very useful. Neither Ward’s nor Average linkage methods have this problem, since they take into account the average distances and so usually indicate differences in between-cluster similarities in this circumstance. These two methods, of course, have other problems in other situations.

With presence-absence data, we used Jaccard’s distance, as described above. When multi-state attributes were used, such as percentages, we usually chose unstandardized city-block distance. A short discussion of the issue of standardization and choice of distance measure is necessary. The justification of the use of unstandardized distance is as follows: standardization equalizes the contribution of each character or attribute (for instance, pottery type) to the total distance. Thus each pottery type would contribute equally to the placement of a site. The occurrence of a rare type would be equally as important as a much higher frequency of a more common type. While there are many situations where standardization is desired (Sneath and Sokol 1973; Matson and True 1974), in most cases it is more intuitively plausible to weight each artifact equally within a site so that types occurring at higher frequencies would be given more weight than those that are less common.
Matson, Lipe, and Haase (Aug. 88) III-65
Therefore, we usually opted for unstandardized distance, although we did use standardized distances at times.

Following common practice, we used city-block or Manhattan distance (average absolute distances), rather than an average Euclidean or traditional taxonomic distance when using multi-state characters (Sneath and Sokal 1973:125). The argument here is that any given difference in frequency should contribute the same amount to the distance measure regardless of whether this difference results from several moderate differences between cases, or from a mix of great and small differences.

For example, if we had two sites, one with 20% red ware and 40% white ware and the other 30% red ware and 30% white ware, the absolute difference would be 20% (10% + 10%), which results in a distance of .20 according to the city block criterion. If we had a third site with 22% red ware and 58% white ware its dissimilarity with the first site would also be 20% (2% + 18%). If we used the Euclidean criterion (i.e., c$^2$=a$^2$+b$^2$) we would find that the first comparison would give us a distance of 14.14, (sq. root of 200, 200=10$^2$ +10$^2$) but the second would give us 18.11 (sq. root of 324, 324=18$^2$+2$^2$). Thus in the cases, a difference of 20% gives us two different Euclidian distances. This does not agree well with the idea of making each potsherd equal, while the unstandardized city block distance does. Interestingly enough, the latter measure is equivalent to the long established favorite in archaeology, the Brainerd-Robinson coefficient (Robinson 1951).
similarity-distance measure are the most important decisions in a project such as this one, because most data reduction techniques will give similar results if the same similarity matrix is used. Even so, this choice is important.

Besides the choice of coefficient, and of clustering method, a third group of decisions must be made about the scaling procedures to be employed. Scaling can use the same distance matrix as a cluster analysis, but rather than producing groups, "dimensions", "factors" and "configurations" result. Here the choices from the classic principal components analysis bases variants, and the newer techniques, both metric and non-metric multidimensional scaling.

As stated earlier, we usually used Torgerson's <metric> scaling procedures (this is the same as Gower's principal coordinates analysis, (Gower 1967)) rather than the nonmetric procedures of Kruskal or Guttman-Lingoes (Kendall 1971, Matson and True 1974, LeBlanc 1975, Drennan 1976) which are more commonly used in archaeology. The reasons for the choice of metric multidimensional scaling (Torgerson 1958) are that when the input matrix is metric, metric scaling gives unique solutions, metric scaling solutions are invariant under changes in dimensionality and metric scaling gives faster solutions and thus has less computational expense (Matson and True 1974).

One of the advantages of nonmetric scaling is that it uses only the rank order of the similarity matrix and thus
Matson, Lipe, and Haase (Aug. 88) III-67 makes no assumptions about the metricity of the similarity or distance coefficients. In our case, city block and Jaccard's distances are known to be metric and this was tested each time by running the distance matrices through the triangle inequality test in our scaling program (Sneath and Sokol 1973:120-21, Anderberg 1973:99-100). In our experience the results to the two scaling techniques are usually similar even under moderate violations of the triangle equality (see also Sneath and Sokol 1973:248-9).

One of the problems with nonmetric multidimensional scaling is that different results do occur if different initial configurations are used. Usually these results do not differ by very much, but this is something that should be checked. Some consensus has developed that a metrical solution is a good initial configuration, as used in TORSCALE (Young 1967) and now in KYST (Kruskal 1974). Even Guttman and Lingoe's SSA uses a technique which can be equivalent to metric scaling for initial configurations. In such situations it has been our experience that the more important dimensions of the nonmetric solutions conform closely to the metric input initial configurations, particularly if the similarity matrix is metric (see for example, Matson and True 1974). If this is the case, there is little reason to go beyond the "initial" configuration, i.e., metric scaling. This fact, plus the uniqueness of its solutions, made metric scaling preferable to nonmetric for most of our use.

One strength of nonmetric scaling, however, is that it
Matson, Lipe, and Haase (Aug. 88) III-68 tends to "solve" input matrices in a space of fewer dimensions than do competing techniques. In many situations, this is a valuable asset, but it is coupled with the disadvantage that the order of the variables on each dimension tends to change if there is a change in the number of dimensions required for a solution. Thus the first dimension in a two dimensional solution may differ from the first dimension in a four dimensional solution. Usually these differences are not too great if the "fit" is good in both solutions. In the noniterative metric scaling procedure, on the other hand the dimensions are "extracted" one by one and do not change in different dimensional solutions so that this problem of changing dimensions does not occur.

Although in this day of low CPU costs differences in computational time are not as important as they were a few years ago, metric scaling offers substantial savings. Not only would metric scaling be carried out anyway as the preferred (or at least an alternate) initial configuration in nonmetric scaling, but the run times are much lower as well as the number of needed runs.

A final series of related advantages of metric scaling is that the importance of the dimensions is known; they come out in order of importance and at least two measures of relative importance are available for the dimensions. Approaches to obtaining equivalent information are available on some, but not all, nonmetric packages and the success of these attempts is variable.
The main advantage of non-metric scaling is that the measures of similarity do not need to be metric. Now that more is known about similarity measures this is not the advantage it appeared 15 years ago. The similarity measures used here are usually metric and do not suffer in comparison with non-metric measures. The common case where non-metric measures are to be preferred is in correlation situations with multistate attributes, and two options are available in this circumstance. The first is to use nonmetric scaling and the second is to use principal components analysis which is appropriate in this situation and very closely related to metric scaling (Gower 1967, Van de Geer 1971:17-23).

The application of metric scaling to our first example data matrix results in Figure III-5. This figure illustrates the two separate clusters found by cluster analysis. In addition to this we can see a trend on the vertical axis. Site A, which only had plain gray pottery, is at the bottom, and sites C and D, which have painted wares but no plain gray, are at the top. The main difference between sites can thus be interpreted as one of time, or as a seriation. The amount of squared distance from the centroid (percent of trace) accounted for by the first two dimensions is 82 percent and 18 percent. The percent of total interpoint distances accounted for by each dimension (Matson and True 1974)—another measure of importance—is 83 percent and 17 percent. The configuration of points is thus solved in two dimensions with the first one clearly the more important, and related to change.
In a number of cases not all sites will have enough variables to allow them to be clustered reliably. In this situation we will define the groups by cluster analysis of the sites on which we do have sufficient information and use multivariate discriminant analysis or inspection to classify the remaining low information sites. This procedure will be discussed in more detail when we need to use it.

Much of the grouping of sites was based on their artifact assemblages. The architectural information was usually used to help interpret the analysis of the artifacts and to set up comparable groups. While in some cases the architectural information is clear cut, in others it is ambiguous and, in yet others, it is missing. Obviously in the last two cases the architecture cannot be easily used to define classes. For example, jacal structures have a much lower archaeological visibility than masonry ones, yet both may have been used for similar purposes. From time to time, however, we will override the analysis of artifactual material where it runs counter to clear-cut architectural information. This frequently occurs where only small surface collections existed. Since surface evidence of features is best preserved in the absence of erosion, and large surface collections are more likely to be made where erosion has been extensive. There is some tendency for large collections to have uninterpretable features and well preserved features to be associated with modest collections.

While the architectural information can be less ambiguous
Matson, Lipe, and Haase (Aug. 88) III-71 than the artifactual, the artifactual information occurs much more frequently and abundantly.

In most of the previous discussion we have been discussing the techniques as being used to group sites, or Q-mode analysis. Since much of our work involves interpreting differences between sites in terms of activities it is also important to discover what artifactual or other variables co-occur together. Analyzing variables in this fashion is referred to as R-mode or R-type analysis.

Since the relationships between variables and sites determine both the R-mode and Q-mode results, both kinds of analyses often reflect the same order. Yet often the understanding of either kind of analysis is enhanced by the other. Because of our interest in activities occurring on sites, R-mode analysis to define functional associations of artifacts is obviously important.

The use of multivariate techniques in archaeology to find "tool kits" has a short but active history. Binford and Binford(1966) initiated this approach, using factor analysis and Middle Paleolithic artifacts. Binford's students, Hill (1968) and Longacre (1970) used similar approaches to try to find groups of ceramic design elements that were associated together. Many others have since followed suit. Our own work has mainly been with surface collections (Matson 1971; 1974a; Matson, Ham and Bunyan 1984), but one attempt used excavated materials (Matson 1976). Most workers have used relative frequency data, but our experience prior to the Cedar Mesa project was primarily with presence-absence data.
Matson, Lipe, and Haase (Aug. 88) III-72 has called into question an implicit assumption of R-mode analysis (1976:15)—that tools found together were used together. Schiffer describes a number of processes that result in artifacts entering archaeological context at locations other than the places they were used or stored. Binford (1973) has developed an idea he calls "curation" to explain why some tools would be kept after use and others discarded, and Schiffer (1972) has suggested that the degree of difference between location of discard and of use is related to site size and intensity of occupation. Binford points out that curation can take place between sites as well.

We argue that curation between Cedar Mesa sites is rarely going to influence our R-mode analyses and that by using sites as provenience units Schiffer’s objections can be circumvented. For site to site curation to introduce major errors into our R-mode analysis a majority of artifacts would have to be discarded on sites where they are not used. The characteristics of the lithic tools and ceramics we use for these analyses makes this possibility quite unlikely. The ceramics are all in the form of potsherds, most of which were presumably discarded when vessels broke. Although individual lithic items (e.g., well-made bifaces) may have been curated, the bulk of our lithics assemblages consists of broken tools that seem likely to have been left or discarded at the site when they were used.

It is difficult to imagine people carrying from site to
Matson, Lipe, and Haase (Aug. 88) III-73 site and discarding on sites where they were not used such items as broken hammerstones, broken retouched flakes, biface fragments, millingstones, manos, mano fragments, and debitage. It is also apparent that for our data, amount of fragments and complete objects are generally correlated, which suggests that this scenario is unlikely for most complete objects as well as broken ones.

For these reasons we argue that using sites as provenience units avoids the potential problems raised by Binford and Schiffer, a position that Schiffer appears to agree with (1976:69). This does not mean that we agree with Schiffer that R-mode analysis is untenable because of discard factors if it is carried out on multiple proveniences within a single site. Following Beckner (1958), we do think if there are two ways of reaching a goal, each equally practical and that one uses less assumptions, then that one should be used. Consequently we have made whole-site assemblages our basic units of analysis in most cases.

An additional problem exists, that has really been brought to the attention of archaeologists since the Cedar Mesa Project was initiated, that of palimpsest, or sites being the result of many visits, over a long period of time, for many different purposes. We did not develop any special method to deal with this problem, but do not think that it is very serious for Cedar Mesa for a variety of reasons. Perhaps most importantly, most of the Cedar Mesa sites were occupied for only short periods of times; in the last four
Matson, Lipe, and Haase (Aug. 88) III-74 chapters we argue that most habitation sites were occupied for less than 15 years. If this is so, then for habitation sites, the palimpsest problem is unimportant. It may be that such sites were also used for other, less intensive activities later, and that these are "swamped" by the residential activities. We will also present arguments that suggest that many of the Basketmaker III and Pueblo II/III limited activity sites are "field stations" which also had similar use durations. The palimpsest problem appears to be most significant for limited activity Basketmaker II sites, and secondary, for re-use of later "field stations". Our procedure, then, was to be aware of the possibilities, rather than to use any particular method to cope with this potential problem, but one that is most likely of only minor importance in this case.

<Summary>

This chapter introduces the major methodological problems and the corresponding solutions developed by the Cedar Mesa Project. Since the Cedar Mesa Project is based on survey sampling, many of the problems and solutions are sampling in nature. For Cedar Mesa we divided the mesa into 20 drainages and sampled three from the north (Upper Grand Gulch, Bullet and North Road) and two from the south-southwest (Hardscrabble and West Johns). We argue that the drainages are valid natural units which have within them most of the environmental variation that exists within the mesa as a whole. The drainages were subsampled by 400
Matson, Lipe, and Haase (Aug. 88) III-75 meter quadrats, at a seven percent rate, so that the individual drainages had from 9 to 22 quadrats collected.

The drainage canyons were a special but rather small (in area) environment. We developed a special sampling scheme for them, by first inventoring all the sites the drainage canyons and then subsampling six different classes of sites. For both quadrats and the sampled canyon sites, the standard procedures were to fill out cultural, physiographic and botanical forms, as well as to grid the site into mapping units. The mapping units were usually either 50 or 20 meters on a side and were used to plot the location of artifact locations, features and trees. With four exceptions, all artifactual material was collected, usually located with a precision of about one meter on the mapping units. Concentrations of artifacts and architectural features were mapped at larger scales, and collected usually with greater precision.

For analytical purposes, the main assumption was that the environmental trends visible today on Cedar Mesa were there in the past. The main exception to this assumption is the arroyo cutting, which has been episodic in the last two or three thousand years. While the local paleoenvironmental information is not detailed enough to specify items such as length of growing seasons, it does indicate an environment similar to today in the recent past.

Because of the non-normal distribution of most quadrat data, we opted to use binomial confidence intervals for the quadrat parameters. Evaluations using simulated data
Matson, Lipe, and Haase (Aug. 88) III-76 demonstrate the superiority of these techniques for
distributions with shapes like that found on Cedar Mesa.
Similarly, we elected to use the powerful Wilcoxon family of
tests for analysis of variance, a family that is not
appreciably inferior in power to the normal t-f tests with
normally distributed data, and incomparably better with many
other shapes.

The same techniques used to evaluate different confidence
intervals were used to evaluate the sampling design. While
both smaller and larger quadrat sizes were found to be more
efficient for certain variables, the 400 meter quadrat size
appears to be a good compromise. While the sample size
appeared to be very useful with the preliminary analysis, we
will report cases later where it is insufficient.

The data reduction techniques used are those of
clustering and scaling with Farthest Neighbor clustering and
Torgerson's metric multidimensional scaling being the
techniques of choice. Besides reviewing the reasons for
choosing these examples of their use are give. The reasons
for using as distance coefficients Jaccard's distance--for
presence-absence data--and city-block distance--for
multi-state data--are presented along with examples.

The chapter concludes with short discussions of
assumptions about site formation processes and their
implications for the analysis of Cedar Mesa data. A way
around one critique of R-mode analysis is presented by using
sites as the unit of association, while the palimpsest
problem is only reviewed.
Figure III-1. Setting of Springs with respect to Canyons and Mesa-top.
Figure III-2. Cedar Mesa Sample. Sampled drainages and quadrats.
Figure III-3. “Typical” Site map. A) Large scale site map, B) Feature map.
The cultural notes included with the site/feature maps should include a description of the nature and distribution of probable structural remains and occupational evidence such as midden, ash fill, etc. This would include size, shape, and volume of structural remains (e.g. slab scatters) and amount and location of disturbances, if any.

a. Surficial observations of fill
   - Clean soil
   - Sandstone chunks
   - Ashy soil
   - Charcoal (burned)
   - Limestone chunks
   - Indeterminate

b. Inferential characteristics of structure
   - Masonry: N/A
     - Type: Dry, Mortared, Indeterminate
     - Preparation of stones: Unshaped, Scabbled, Well-shaped, Other
     - Slab construction: 
     - Type of shaping: Unshaped, Scabbled, Well-shaped, Other, Indeterminate
     - Jadal construction: N/A
     - Pits: N/A
       - Roofed (indicate cribbing, beams, vigs, etc.)
       - Roofless
       - Indeterminate
     - Inferred size and shape of structure: slab-lined, hearth, etc.
     - Kiva: N/A
       - Inferred function of structure: Storage, Windbreak, or other temporarily used, room type structure, Indeterminate
     - Other: N/A
       - Kiva
       - Habitation
       - Indeterminate
     - Retaining Wall
     - Defensive Wall
     - Water
     - Control Device (specify): Other (specify)
     - Indeterminate

c. Non-structural feature
   - Burial
   - Midden
   - Ashy spot
   - Chipping station
   - Burned sandstone and limestone cluster
   - Other (specify)

 d. Results of subsurface investigation
    - Location of test unit:
      - Flagged hearth and ash spot
    - Fill characteristics (note the maximum depth of occupational fill observed): 
      - Hearth - 0.75m deep
      - Ashy spot - 0.5m deep
      - 1.0m deep
      - Other (specify) cultural material
      - Indeterminate

 e. Prelim estimate of date, cultural affiliation (specify basis of estimate)
    - Presence of pottery and burned debris
    - Other (specify)
    - Other cultural material

 f. Photos taken (specify subject, film and exposure no.)

 g. Recommendations for further work (e.g., test pits, flotation, subsistence information, etc.)

Figure III-4. Cultural Form example.
Figure III-4. Cultural Form example.

Figure III-5. Botanical Form Example.
Figure III-6. Physiographic Form example.

Matson, Lipe, and Haase  Chapter III Figures, Feb. 12, 2010 – 6
Figure III-7a. Canyon Survey Form Example.
Figure III-7b Canyon survey sketch map example.
<table>
<thead>
<tr>
<th>I. Watershed area</th>
<th>II. Water source number</th>
<th>III. Date of survey</th>
<th>IV. Ease of access to source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bullit</strong></td>
<td></td>
<td></td>
<td><strong>Easily accessible</strong> from rim, long distance from camp,</td>
</tr>
<tr>
<td><strong>Upper Grand</strong></td>
<td></td>
<td></td>
<td><strong>Moderately difficult of access</strong></td>
</tr>
<tr>
<td><strong>Hardscrabble</strong></td>
<td></td>
<td></td>
<td><strong>Very difficult of access</strong></td>
</tr>
<tr>
<td><strong>West Johns</strong></td>
<td></td>
<td></td>
<td><strong>Accessible only to wildlife</strong></td>
</tr>
<tr>
<td><strong>Upper Road</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

V. Type of water source (add descriptive comments at right, if necessary)
- Seep
- Flowing spring
- Pothole or tank
- Other (describe at right)

VI. Variability of source. Water is available:
- All seasons
- Fall
- Spring
- Winter
- Summer
- Wet years only
- Normal years
- Even in dry years

VII. Productivity of source.
- For seeps and springs, estimate available water, in gallons per day.
- Normal year, season of max. yield if source cleaned
- Normal year, season of min. yield if source uncleaned
- Comments:
- a 10 gal standing pool with a very slight runoff under overhanging a very head of red cacti

VIII. Type of physiographic situation
- Canyon head (point of entrenchment)
- Canyon bottom, steep upper part
- Canyon bottom, gentle lower part
- Masa top
- Other

IX. Type of geological situation
- Bedding plane or joint in bedrock
- Contact between pervious and impervious layers (specify formations involved if different)
- Talus or landslide deposits
- Dune sand
- Alluvium
- Colluvium

X. Vegetation. Indicate plants noted at springs or seeps. Continue on back if necessary
- Red & green algal

XI. Animal sign noted at water source. Continue on back if necessary
- **Elephant** Chipmunk, Water Bugs

---

Figure III-8. Water Resources Form Example

Matson, Lipe, and Haase Chapter III Figures, Feb. 12, 2010 – 9
Figure III-9. Cluster Analysis dendrogram of Table III-8.

Figure III-10. Cluster analysis dendrogram of Table III-10.

Figure III-11. Multidimensional Scaling of Table III-8.
## Chapter III Tables

<table>
<thead>
<tr>
<th>Table No.</th>
<th>Table Heading</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table III-1</td>
<td>Drainage Canyon Sampling Summary</td>
<td>2</td>
</tr>
<tr>
<td>Table III-2</td>
<td>Comparison of “Normal” Confidence Intervals on Mean with “Binomial” Intervals on Median.</td>
<td>3</td>
</tr>
<tr>
<td>Table III-3</td>
<td>Bullet Drainage Summary Site Statistics for Quadrats</td>
<td>3</td>
</tr>
<tr>
<td>Table III-4</td>
<td>Comparison of Simulation of 400 Meter Quadrat Samples of “Unit” Type Sites from Bullet With 800 Meter Quadrat</td>
<td>4</td>
</tr>
<tr>
<td>Table III-5</td>
<td>Upper Grand Gulch 400 Meter Quadrat “Unit” Type Site Simulation</td>
<td>4</td>
</tr>
<tr>
<td>Table III-6</td>
<td>Comparison of Simulation of 200 and 400 Meter Quadrat Samples</td>
<td>5</td>
</tr>
<tr>
<td>Table III-7</td>
<td>Hypothetical Data Array: A,B,C, and D sites</td>
<td>5</td>
</tr>
<tr>
<td>Table III-8</td>
<td>Matrix of Jaccard’s Distance</td>
<td>5</td>
</tr>
<tr>
<td>Table III-9</td>
<td>Recalculated Distance Matrix</td>
<td>6</td>
</tr>
<tr>
<td>Table III-10</td>
<td>Hypothetical Data Array: A,B,C,D,E and F sites</td>
<td>6</td>
</tr>
</tbody>
</table>
### Table III-1

Drainage Canyon Sampling Summary  
(Not including quadrat canyon sites)

<table>
<thead>
<tr>
<th>Drainage</th>
<th>Site Class</th>
<th>Inventoried</th>
<th>Sampled</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM II</td>
<td></td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>BM III</td>
<td></td>
<td>3</td>
<td>4*</td>
</tr>
<tr>
<td>Upper</td>
<td>PII/III Habitation</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Grand</td>
<td>PII/III Non-Habitation</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>Gulch</td>
<td>Specials</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BM II</td>
<td></td>
<td>29</td>
<td>7</td>
</tr>
<tr>
<td>BM III</td>
<td></td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Bullet</td>
<td>PII/III Habitation</td>
<td>26</td>
<td>5</td>
</tr>
<tr>
<td>Canyon</td>
<td>PII/III Non-Habitation</td>
<td>31</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Special</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>BM II</td>
<td></td>
<td>17</td>
<td>6</td>
</tr>
<tr>
<td>BM III</td>
<td></td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>North</td>
<td>PII/III Habitation</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Road</td>
<td>PII/III Non-Habitation</td>
<td>25</td>
<td>8</td>
</tr>
<tr>
<td>Canyon</td>
<td>Special</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>BM II</td>
<td></td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>BM III</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hard-scrabble</td>
<td>PII/III Non-Habitation</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Canyon</td>
<td>Special</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BM II</td>
<td></td>
<td>25</td>
<td>7</td>
</tr>
<tr>
<td>BM III</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>West</td>
<td>PII/III Habitation</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>John’s</td>
<td>PII/III Non-Habitation</td>
<td>24</td>
<td>5</td>
</tr>
<tr>
<td>Canyon</td>
<td>Special</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>BM II</td>
<td></td>
<td>84</td>
<td>33</td>
</tr>
<tr>
<td>BM III</td>
<td></td>
<td>14</td>
<td>15*</td>
</tr>
<tr>
<td>Totals</td>
<td>PII/III Habitation</td>
<td>71</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>PII/III Non-Habitation</td>
<td>104</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Special</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>291</td>
<td>115</td>
</tr>
</tbody>
</table>

* UG C-18-1 was inventoried as a P II/III habitation site, but when it was sampled and collected it was found to be a BM III site. Another PII/III habitation site was then selected for collection.
Table III-2
Comparison of “Normal” Confidence Intervals on Mean with “Binomial” Intervals on Median.

Simulated data, sample size is 10, 100 samples.

<table>
<thead>
<tr>
<th></th>
<th>95% “Student” Intervals</th>
<th>98% “Binomial” Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means within</td>
<td>85</td>
<td>Medians within</td>
</tr>
<tr>
<td>Means outside</td>
<td>15</td>
<td>Medians outside</td>
</tr>
<tr>
<td>Means low side</td>
<td>15</td>
<td>Medians tied low side</td>
</tr>
<tr>
<td>Means high side</td>
<td>0</td>
<td>Medians tied high side</td>
</tr>
</tbody>
</table>

Simulated “Cedar Mesa” data, 400 meter quadrats (see Table III-3), sample size is 21, 100 samples.

<table>
<thead>
<tr>
<th></th>
<th>95% “Student” Intervals</th>
<th>98% “Binomial” Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means within</td>
<td>83</td>
<td>Medians within</td>
</tr>
<tr>
<td>Means outside</td>
<td>17</td>
<td>Medians outside</td>
</tr>
<tr>
<td>Means high side</td>
<td>17</td>
<td>Medians high side</td>
</tr>
<tr>
<td>Means low side</td>
<td>0</td>
<td>Medians low side</td>
</tr>
</tbody>
</table>

* These are medians on the boundary of the confidence intervals.

Table III-3
Bullet Drainage Summary
Site Statistics for Quadrats(n=21).

<table>
<thead>
<tr>
<th></th>
<th>BM II</th>
<th>BM III</th>
<th>P II-III</th>
<th>?</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site Medians</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Site Interquartiles</td>
<td>0-2</td>
<td>0-2</td>
<td>1-4</td>
<td>0-2</td>
<td>4-9</td>
</tr>
<tr>
<td>Median 90% Binomial Confidence intervals</td>
<td>1-2</td>
<td>0-1</td>
<td>2-4</td>
<td>0-2</td>
<td>4-8</td>
</tr>
<tr>
<td>Site Means</td>
<td>1.67</td>
<td>.95</td>
<td>2.90</td>
<td>1.43</td>
<td>6.81</td>
</tr>
<tr>
<td>Total artifact Medians</td>
<td>147</td>
<td>50</td>
<td>525</td>
<td>4</td>
<td>1230</td>
</tr>
<tr>
<td>Total artifact Interquartiles</td>
<td>0-508</td>
<td>0-312</td>
<td>382-1503</td>
<td>0-17</td>
<td>721-2110</td>
</tr>
<tr>
<td>Median 90% Binomial Confidence Intervals</td>
<td>41-228</td>
<td>0-278</td>
<td>390-1132</td>
<td>0-10</td>
<td>956-1674</td>
</tr>
<tr>
<td>Artifact Means</td>
<td>340.8</td>
<td>235.2</td>
<td>929.3</td>
<td>9.0</td>
<td>1514.3</td>
</tr>
</tbody>
</table>

Matson, Lipe, and Haase Chapter III Tables, August 13, 2010—3
### Table III-4
Comparison of Simulation of 400 Meter Quadrat Samples of “Unit” Type Sites from Bullet With 800 Meter Quadrat Sampling Simulation
(20 samples, each of 21 quadrats)

<table>
<thead>
<tr>
<th></th>
<th>800 Meter</th>
<th>400 Meter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample mean (of means)</td>
<td>.122</td>
<td>.129</td>
</tr>
<tr>
<td>Sample median (of means)</td>
<td>.12</td>
<td>.14</td>
</tr>
<tr>
<td>Interquartile(6, 15) (of means)</td>
<td>.95, .143</td>
<td>.10, .14</td>
</tr>
<tr>
<td>Low (of means)</td>
<td>.067</td>
<td>0.0</td>
</tr>
<tr>
<td>High (of means)</td>
<td>.17</td>
<td>.29</td>
</tr>
</tbody>
</table>

Note: Population mean is .12 per 160,000 sq. meters.

### Table III-5
Upper Grand Gulch 400 Meter Quadrat “Unit” Type Site Simulation
(100 samples, each of 9 quadrats)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample mean (of means)</td>
<td>.445</td>
</tr>
<tr>
<td>Sample median (of means)</td>
<td>.375 (.375, .500)</td>
</tr>
<tr>
<td>Interquartile (25, 76)</td>
<td>.25, .75</td>
</tr>
<tr>
<td>Low (of means)</td>
<td>0.0{6}</td>
</tr>
<tr>
<td>High (of means)</td>
<td>1.0{1}</td>
</tr>
</tbody>
</table>

Note: Population mean is .46. ()s enclose 95% confidence intervals of median. {}s indicate number of samples with that value.
### Table III-6
Comparison of Simulation of 7% 400 Meter Quadrat Samples of Pueblo II-II Material with 2.4% 200 Meter Quadrat Samples (100 Samples each)

<table>
<thead>
<tr>
<th>400 Meter</th>
<th>200 Meter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number in sample</td>
<td>21</td>
</tr>
<tr>
<td>Pop. Mean (per 160,000 sq. m.)</td>
<td>1202.2</td>
</tr>
<tr>
<td>Population median</td>
<td>970</td>
</tr>
<tr>
<td>Sample mean (of means)</td>
<td>1172.7</td>
</tr>
<tr>
<td>Sample median (of means)</td>
<td>1130</td>
</tr>
<tr>
<td>Sample interquartile (of means)</td>
<td>969, 1365</td>
</tr>
<tr>
<td>Sample low (of means)</td>
<td>676</td>
</tr>
<tr>
<td>Sample high (of means)</td>
<td>1957</td>
</tr>
</tbody>
</table>

### Table III-7
Hypothetical Data Array: A,B,C, and D are sites; + and - indicate presence or absence.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Plain Gray</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>2.</td>
<td>Corrugated Gray</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>3.</td>
<td>Black-on-White</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4.</td>
<td>Red-on-Orange</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

### Table III-8
 Matrix of Jaccard’s Distance

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.50</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1.0</td>
<td>0.75</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1.0</td>
<td>1.0</td>
<td>0.33</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Matson, Lipe, and Haase Chapter III Tables, August 13, 2010—5
Table III-9
Recalculated Distance Matrix

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C,D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>.50</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>C,D</td>
<td>1.0</td>
<td>1.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table III-10
Hypothetical Data Array: A,B,C,D,E and F are sites; + and - indicate presence or absence.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Plain Gray</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2.</td>
<td>Corrugated Gray</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>3.</td>
<td>Black on White</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>4.</td>
<td>Red on Orange</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>