DOLORES ARCHAEOLOGICAL PROGRAM TECHNICAL REPORTS

Report Number: DAP-079

Explaining variability among artifact collections

by

Timothy A. Kohler, Carl J. Phagan, and Eric Blinman

ANASAZI HERITAGE CENTER
LIBRARY

Prepared for
Cultural Resources Mitigation Program: Dolores Project
Bureau of Reclamation, Upper Colorado Region
Contract No. 8-07-40-S0562

Under the supervision of
David A. Breternitz, Senior Principal Investigator

Final Submission

4 August 1983

TABLE OF CONTENTS

	Page Number
LIST OF FIGURES	ii
LIST OF TABLES	iii
ABSTRACT	iv
INTRODUCTION	1
DEFINITION OF THE PROBLEM	4
ANALYSIS OF COLLECTION BIAS BY SITES	7
ANALYSIS BY TEMPORAL-SPATIAL UNITS	13
DISCUSSION AND CONCLUSIONS	24
REFERENCES CITED	26

LIST OF TABLES

Page Number

Table 1. Summary of variables employed in the site analysis 8

Table 2. Spatial/temporal elements used in analysis 15

Table 3. Substantive variables used in the analysis by element. . . . 16

Table 4. Collection variables used in the analysis by element 17

Table 5. Relationships between major temporal, spatial, and functional classes and the substantive variables from table 3. . 19

Table 6. Relationships between temporal, spatial, and functional classes and the collection variables from table 4. 20

Table 7. Correlations between collection variables and substantive variables 21

LIST OF FIGURES

Page Number

Figure 1. Formation of intercollection variability. 2

ABSTRACT

Not all variability in archaeological collections is due to phenomena in which archaeologists are interested. Significant variability among collections from sites and portions of sites excavated by the Dolores Archaeological Program in southwestern Colorado can be attributed to differences in mode of collection, types of study units, and deposits encountered in sampling.

INTRODUCTION

> The subject of archeology is behavioral and cultural variability - the similarities and differences in behavior at specific temporal and spatial loci. Its object is to explain these similarities and differences (Plog 1974:13).

Explanation of variability within and between collections of archaeological materials is a basic activity of the archaeologist. In most cases, variability is assumed to be related to prehistoric cultural differences, and examples of interpretations cover considerable ground. Perceived variability has led to interpretations of ethnic affiliation (Bordes 1972:146-159), spatial segregation of activities (Binford and Binford 1966), postmarital residence rules (Longacre 1970), economic specialization (Rice 1981), social differentiation (Otto 1975), dichotomies between sacred and secular (Sears 1973), degrees of interaction between populations (Englebrecht 1978), environmental influences (Arnold 1978), intensity of occupation (Cook 1972), duration of occupation (Schiffer 1975), and temporal differences among assemblages (Ford 1952). In some cases, variability has also been attributed to circumstances surrounding the prehistoric abandonment of site locations (Kane and Robinson 1980), later cultural modification of earlier deposits, and noncultural postdepositional processes (Schiffer 1976:15). Variation in the history of the deposits resulting from any of these factors can alter the set of materials available for study in the archaeological context (fig. 1).

Recently archaeologists have become aware that the nature of the <u>sample</u> of materials drawn from an archaeological context may affect interpretations based on those materials. Differences between judgmental and probabilistic samples have been recognized (Kohler and Gross 1981), and Nance (1981) has discussed how the extent of excavation determines the

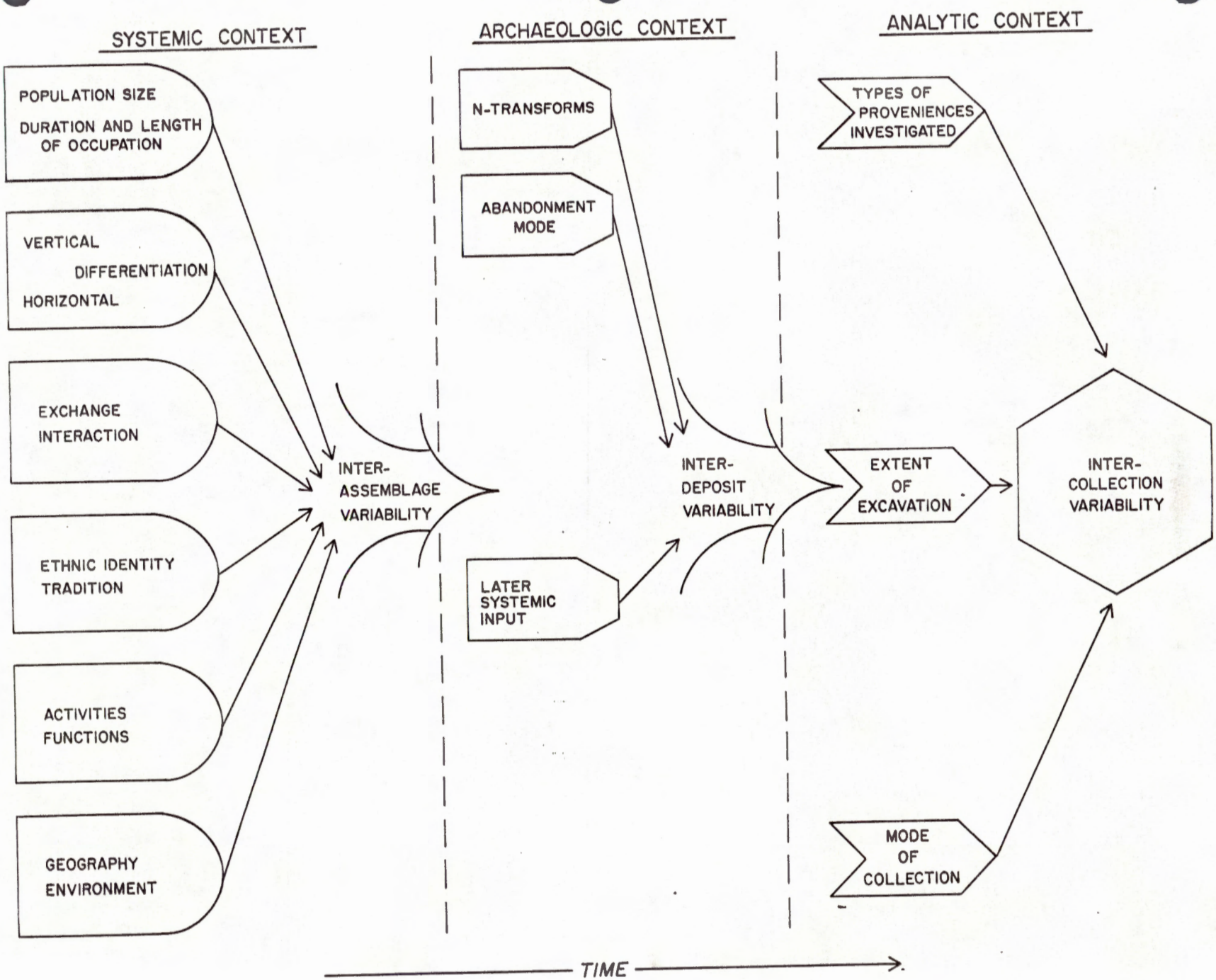

Figure 1. Formation of intercollection variability.

likelihood that rare artifact types will be included in a sample. Under the heading Analytic Context in figure 1 we have identified several factors affecting the composition of collections recovered from archaeological deposits. We will reserve the term "assemblage" for materials in use in a systemic context; "deposits" for the materials and their spatial relations in an archaeological context; and "collections" for materials taken from the archaeological context. The purpose of this paper is to examine the effects of archaeological recovery techniques and types of proveniences investigated on collections from sites, and on collections from spatially and temporally defined portions of sites in the Dolores River valley.

DEFINITION OF THE PROBLEM

Variability within and between archaeological collections has been used for a variety of purposes by the DAP (Dolores Archaeological Program). Date estimates for proveniences have been made on the basis of relative frequencies of pottery types (Blinman 1982) and on variation in utility ware neckband heights (Blinman 1981); sites have been classified into a functional typology based on total artifact collections (Schlanger and Harden 1983); and lithic profiles have been used to differentiate Archaic and Anasazi lithic scatters (Phagan 1981). Such studies require various assumptions about how data from collections represent particular systemic contexts. Some of these studies, and many archaeological hypotheses in general, can be addressed using data consisting of diagnostic items, spatial relationships, or inferred functions. Other studies require data from collections that adequately represent systemic assemblages. These latter studies appear to be particularly sensitive to variation introduced by the archaeologist.

From the beginning of the DAP excavation program in 1978 there were important factors introducing variability among collections of materials. Some of these are in part systemic; sites are dissimilar in setting, structure, size, depth and density of deposits, and content. Others are nonsystemic. Two steps were taken by the DAP to achieve as much collection consistency as possible:

1. All excavation was conducted according to a standard field manual that was first drafted in 1978, and which was later revised and expanded (Kane, Hewitt et al. 1981). The manual contains guidelines for surface collection and excavation procedures, including suggestions as to

when screening, troweling, excavation by heavy equipment, and so forth, may be appropriate.

2. Field forms were established that required description of the nature of the deposit where materials originate and how collections were made. For purposes of studying intersite comparability of data the most important of these forms is the Field Provenience Description Form. This form is filled out for each excavation lot, and includes information on the following.

   a. Study unit type (e.g., surface room, kiva, midden; 1- by 1-m square, trench).

   b. Fill/assemblage position (relating the position of excavated fills to observable cultural units: roof fall, wall fall, floor surfaces, foundations).

   c. Fill/assemblage type (differentiating types of cultural deposits, e.g., primary, secondary, and de facto refuse; materials in apparent use association; deposits with a noncultural or mixed origin).

   d. Mode of collection (e.g., various-sized screens, nonintensive and intensive surface collection, trowel, trowel and shovel, backhoe).

Coding these attributes for each provenience provides the DAP with one means for determining the extent of intercollection variability due to factors in the analytic context.

These field procedures were augmented in 1979 by the initiation of a probability sampling program for selected sites. In this program, collection techniques were standardized, with all matrix screened through one-quarter-inch (6.44 mm) mesh and all observed cultural materials

recovered. However, these probability samples were designed to provide representative samples of populations that could be spatially designated prior to excavation. Correspondence between prior designated populations (for example, all materials in a site) and ideal units of interpretation (elements and households in DAP terminology) is rare, but these samples do have value in helping discover collection biases in sites not so sampled (Kohler and Gross 1981).

Despite the standardization imposed by the field manual, excavation forms, and the probability sampling program, research strategies inevitably differ among sites to a certain extent. Differences arise in part from perceived research potential in relation to the Dolores Archaeological Program General Research Design (Kane, Lipe et al. 1981), time available for investigation, weather, and accessibility by heavy equipment. Complete standardization is neither possible nor desirable. But how much variability is introduced by these factors? Is it small compared with variability due to systemic factors, suggesting that it can be ignored? Or would failure to control for such factors invalidate across-collection comparisons? In the next two sections, various statistical approaches to this problem are reported.

ANALYSIS OF COLLECTION BIAS BY SITES

A preliminary analysis of potential collection bias was carried out on 66 DAP collections in spring 1982. These collections were aggregated by site, and the sites represented all phases, functions, and sizes known to exist in the project area. These 66 sites were examined in the field with varying degrees of thoroughness, from intensive excavation with considerable horizontal exposure, to sampling using probabilistic techniques, to intensive surface collection. The strategy in this first analysis was to compare the proportions of proveniences collected in various manners with the numbers and weights of various kinds of materials recovered to explore for variation correlated with the collection strategies utilized. Pearsonian correlation coefficients were computed between the "shape" variables of the provenience data and the "size" variables of the collection data. In this exploratory analysis, little attention was paid to ensuring that the assumptions of the product-moment correlation coefficient as a Best Linear Unbiased Estimator (Lewis-Beck 1980:27) were met. Table 1 lists the variables describing the materials in the collections, which were correlated with provenience and mode of collection data. Results are reported below if the associated (and perhaps occasionally inappropriate) significance test indicated that such results could have been achieved less than 1 percent of the time by chance alone, if there were no linear relationship between the variables in question.

Several significant correlations were found:

1. As the proportion of proveniences in the collections representing modern ground surface increased,

   - thick biface weight
   - projectile point count
   - hammerstone weight

   decreased across this sample of 66 sites. About 10 percent of the variation in the values of these variables can be accounted for in this fashion as measured by the Pearson product-moment correlation coefficient. Some of these effects may be explained by activities of modern collectors; the negative correlation between projectile point count and proportion of the total proveniences on the surface suggests projectile points were differentially collected from site surfaces before the DAP collections were made.

Table 1. Summary of variables employed in the site analysis

Ceramics	Nonflaked lithic tools
Bowl sherds	Hammerstones
Jar sherds	Mano fragments, nfs
	One-hand manos
Flaked lithic tools	Two-hand manos
Utilized flakes	Metate fragments, nfs
Cores	Trough metates
Used cores/cobbles*	Slab metates
Thick unifaces	Basin metates
Thin unifaces	Hafted items
Specialized forms†	Ornaments
Thick bifaces	
Thin bifaces	Flaked lithic debitage
Projectile points	

* Includes items referred to as hammerstones.
† Includes forms referred to as spokeshaves, denticulates, burins, gravers, perforaters, drills, and flaked axes.

NOTES: Both counts and weights were used for each variable.
nfs - Not further specified.

2. As the proportion of screened proveniences increased,

   - debitage count and weight
   - core weight
   - thick uniface weight
   - projectile point count and weight

   increased as well. Note that these represent the total weights for a site and not mean weights per artifact type. Screening increases the amounts of almost all materials recovered, but it appears to do so disproportionately for these items. This is not surprising for items small in size such as debitage and projectile points, which may easily be missed in unscreened samples. Ten to 15 percent of the variance in the amounts of these artifacts recovered is correlated with the proportion of proveniences screened in this sample.

3. Where one chooses to excavate within a site is even more important than how one choses to excavate. Thirty-six of the 44 counts and weights in table 1 increased significantly as the proportion of proveniences within structure fills increased. The only variables not significantly affected were

   - thin uniface weight and count
   - mano fragment (nfs) weight
   - metate fragment (nfs) count
   - slab metate count
   - basin metate weight and count
   - nonflaked lithic ornament weight.

   Structure fills often include primary and secondary refuse rich in most artifacts. It seems unlikely, however, that high-input items

such as ornaments would be found in such situations. The small sample size for ornaments (count=102), slab metates (count=49), and basin metates (count=5) may also explain, in part, the failure of these variables to follow this pattern. The materials are all more strongly associated with cultural surfaces than with structure fills (see the following discussion). Ten to 30 percent of the variation in counts and weights of those items not appearing on the preceeding list was associated with variation in the proportion of proveniences from structure fill.

4. Many intersite and intrasite analyses of artifacts focus particularly on materials found in contact with a cultural surface, because their temporal and spatial relations may be more interpretable. In our sample, as the proportion of proveniences in contact with a cultural surface <u>increased</u>,

   - jar weight
   - bowl weight and count
   - utilized flake weight and count
   - used core/cobble tool weight and count
   - thick uniface weight and count
   - thick biface weight and count
   - thin biface weight
   - projectile point weight and count
   - hammerstone weight and count
   - mano fragment (nfs) weight and count
   - one-hand mano count
   - two-hand mano weight and count

- trough metate weight
- hafted item weight and count
- ornament count

all increased as well. Between 10 and 35 percent of the variance in these variables is associated with variance in the proportion of proveniences originating on cultural surfaces.

It is obvious that variability in how sites were excavated and where excavation took place within sites is correlated with, and probably responsible for significant variability in the collections obtained. We note that this is true for materials from the DAP where there has been considerable continuity of excavators and research design, and all excavation has been conducted according to the specifications in a single field manual. It would be all the more true where there is an attempt to compare materials obtained under less controlled conditions.

This first analysis suggests that simply ignoring variability due to collection techniques may promote spurious results for some, especially assemblage related, questions. However, there are several weaknesses in this preliminary analysis. The analysis is based on whole sites, which are rarely used as units for comparison in the DAP although they are common currency in many other reports and analyses. The analysis used proportions of proveniences collected in various manners as a proxy for proportions of materials actually recovered from those proveniences. The variables examined were simple counts and weights rather than synthetic variables which may be central to later analyses. Finally, there was no attempt to consider whether some of the relationships noted here between substantive variables and collection factors might be due in part to complex relationships among collection factors, various spatial, temporal,

and functional distinctions, and substantive variables. For example, the perceived relationship between increased counts and weights of debitage recovered under increasing screening could conceivably result from an unrecognized tendency to screen more proveniences in sites of types or ages having especially high densities of debitage. Simply put, we had discovered correlations but had not demonstrated causes. We set out to remedy as many of these weaknesses as possible in a second analysis examining the effect of collection variability on selected variables and units that would be important in synthetic work.

ANALYSIS BY TEMPORAL-SPATIAL UNITS

The vast majority of sites within the Dolores Valley were occupied between A.D. 600 and 950. These sites represent what is commonly referred to as the Basketmaker III and Pueblo I periods of northern Anasazi prehistory. Intensity of occupation in the valley was high during these periods, and sites are often multicomponent, exhibiting either continuous or discontinuous sequences of occupation. For this reason, temporal and spatial subdivisions of sites have been defined for analytic purposes by the DAP.

The basic temporal unit is the "element," which is defined to represent the material remains of a significant and discrete occupation at a site. "Significant" implies a continuous occupation for periods of perhaps a month or more each year with repetition of this occupation pattern over several years. "Discrete" implies a total duration of use not exceeding the expected lifespan of a household as a cooperating unit, which we take to be about one generation. Elements are usually associated with architectural features and are defined to include both seasonal and permanent habitations. The minimal spatial correlate of the element is one household, but an element may include the remains of multiple households if they are contemporary. Since the dating resolution for elements is often lower than the generation implied by the definition, elements are subdivisions of sites and do not crosscut sites.

Materials are assigned to elements based on field observations and inferences of stratigraphic associations. Assignments are qualified by estimates of the integrity of the materials and by the confidence in the associations. Collections vary in the amount of suspected contamination

or mixing of materials from various elements, and the strength of the inferred association of particular materials with a particular element can be assessed. A collection associated with an element with high confidence and integrity is assumed to represent the unequivocal and uncontaminated material remains of a significant and discrete occupation.

Elements within sites usually can be compared with each other using temporal relationships based on stratigraphy, but comparisons between elements from different sites must rely on date estimates for each element. Tree-ring dated elements are uncommon; therefore most date estimates are based on tree-ring calibrated patterns of ceramic change. Ceramic dating resolution is only adequate for the definition of relatively broad time periods, and these periods are used to group elements for the following analyses (table 2).

In this second analysis we focused on the different and somewhat smaller set of variables in table 3, which we anticipated using in later synthetic work. Each variable in tables 3 and 4 was broken down by period, site type, and a spatial variable that divided the elements in the analysis into five geographic groups (localities). Collection variables were computed separately for each major material category (major headings in table 3) resulting in variables such as proportion of ceramics collected from surface structures, proportion of debitage collected from pitstructures, and so forth. Each of the substantive variables (table 3) was further examined for correlation with the collection variables in table 4. Correlations between substantive and collection variables were measured with Kendall's Tau-b (Blalock 1979:436-439), an ordinal correlation coefficient chosen after an initial univariate screening showed many nonnormal distributions with positive skew for which it was not possible

Table 2. Spatial/temporal elements used in analysis

	Site No.	Excavator	Season	Element	Type	Location
EARLY PERIODS						
Period 1 (A.D. 600-725)						
	5MT2162	Greenwald	1979	2	Ha	S Sagehen
	5MT2198	Hewitt	1979	1	Ha	S Sagehen
	5MT2848	Greenwald	1979	1	Ha	S Sagehen
	5MT2858	Montgomery	1979	2	Ha	N Sagehen
	5MT4545	Brisbin	1979	2	Ha	S Sagehen
	5MT4613	Nelson	1981	1	Ha	Dolores
	5MT4684	Tucker	1980	1	Ha	S Sagehen
	5MT4684	Tucker	1980	2	Ha	S Sagehen
Period 2 (A.D. 725-800)						
	5MT2193	Emerson & Brisbin	1978 & 1979	1	Ha	S Sagehen
	5MT2193	Emerson & Brisbin	1978 & 1979	2	Ha	S Sagehen
	5MT2848	Greenwald	1979	2	Ha	S Sagehen
	5MT2853	Greenwald	1979		Ha	S Sagehen
	5MT2854	Kuckelman	1980	2	Ha	N Sagehen
	5MT4614	Yarnell	1979	2	Ha	S Sagehen
	5MT4640	Hewitt	1979	1	Ha	N Sagehen
	5MT4644	Brisbin	1979 & 1980	1	Ha	N Sagehen
	5MT4644	Brisbin	1979 & 1980	2	Ha	N Sagehen
Period 3 (A.D. 800-840)						
	5MT2161	Sebastian & Hogan	1979 & 1980	1	Se	Grass Mesa
	5MT2192	Yarnell	1978	1	Ha	S Sagehen
	5MT2194	Brown	1979	1	Ha	S Sagehen
	5MT4512	Wilshusen	1979	1	Se	S Sagehen
	5MT4644	Brisbin	1979 & 1980	3	Ha	N Sagehen
	5MT4671	Wilshusen	1979	2	Ha	N Periman
	5MT4671	Yarnell	1979	1	Ha	N Periman
LATE PERIODS						
Period 4 (A.D. 840-880)						
	5MT2161	Sebastian & Hogan	1979 & 1980	2	Ha	Grass Mesa
	5MT2191	Hewitt	1978	1	Se	S Sagehen
	5MT4479	Kleidon	1980	1	Ha	S Sagehen
	5MT4650	Gross	1980	1	Ha	Grass Mesa
	5MT5106	N. Morris	1981	1	Ha	S Sagehen
Period 5 (A.D. 880-910)						
	5MT2205	Kleidon	1979	1	Ha	S Sagehen
Period 7 (A.D. 950 +)						
	5MT2241	M. Morris	1981	1	Ha	S Sagehen

NOTE: Ha - Habitation.
Se - Seasonal site.
S - South.
N - North.

Table 3. Substantive variables used in the analysis by element

Ceramics
- PSM $\geq$ 4*
- Dolores Tract†
- "Late" types§
- Modified items
- Jars

Flaked lithic tools
- Mean weight
- Nonlocal materials
- Utilized flakes
- Projectile points
- Complete or nearly complete
- Well-shaped or stylized

Nonflaked lithic tools
- Mean weight
- Complete or nearly complete
- Hafted items
- Ornaments
- Metates (all subtypes)
- Manos (all subtypes)
- Hammerstones
- Well shaped or stylized
- Resurfaced with light vertical force**
- Resurfaced with medium vertical force**
- Resurfaced with heavy vertical force**

Flaked lithic debitage
- Mean weight
- Cortex††

---

* PSM (production step measure) is a relative index of labor invested into ceramic manufacture (Feinman et al. 1981; Blinman and Wilson 1982).
† Potsherds that are not demonstrably nonlocal on the basis of temper.
§ Mancos Gray, Cortez Black-on-white, Deadman's Black-on-red, slipped white wares, any ceramics with sherd temper, and any corrugated ceramics are considered to be late.
** Wear that is distinguishable as resulting from maintenance, or from maintainance in addition to production or use.
†† Proportion of flakes having any cortex on dorsal surface.

NOTE: Each variable - except mean weight - is the proportion of the frequency of that variable to the total count for its category. For example, the category "Well shaped or stylized" flaked lithic tools is the proportion of the flaked lithic tools so formed to all flaked lithic tools, for each element.

to find an acceptably normal reexpression. Whereas the first analysis had correlated provenience "shape" variables with collection "sizes," the correlations in this analysis are between shape variables.

Table 4. Collection variables used in the analysis by element

	Sites			Elements		
	$\overline{X}$	s	CV	$\overline{X}$	s	CV
Study unit type						
Surface structures	.095	.118	125	.134	.226	168
Pitstructures	.183	.206	112	.341	.366	107
Nonstructures*	.056	.109	194	.094	.202	216
Modern ground surface	.102	.147	143	.038	.094	249
Refuse type						
De facto refuse	.116	.167	144	.224	.305	136
Use association	.016	.056	351	.039	.152	391
Mode of collection						
Screened	.045	.118	260	.224	.311	139
Troweled	.084	.158	187	.321	.352	110

* Outside use areas such as middens, courtyards, plazas, and ramadas.

NOTES: The mean ($\overline{X}$), standard deviation (s), and the coefficient of variation (CV) are compared for the whole site data and the element data. For the data set organized by site, the proportions are the proportions of all proveniences collected. For the elements, the proportions given were calculated for ceramics, but illustrate those for other material classes as well.

$$CV = \frac{s}{\overline{X}} \times 100.$$

Table 4 also summarizes the differences in variability between the "whole site" data set and the data set composed of materials from proveniences assignable with high confidence and integrity to an element. Although the variability in some collection-related variables is lower in the element data, it is higher in others. There is no clear superiority of element over site data in minimizing variability on these collection

variables. There is a clear trend towards a higher proportion of materials from more interpretable proveniences and more intensive collection techniques in the element data.

This analysis - summarized in tables 5, 6, and 7 - shows that several of the collection and substantive variables have a significant relationship with time (as measured either by the 5 periods or by a dichotomization of the sequence into early and late); space (as measured either by the 5 spatial groups or by an upstream/downstream division); and function (either habitation or seasonal).

Table 5 lists significant relationships between the substantive variables from table 3 and these selected temporal, spatial, and functional divisions. Table 6 presents similar information for the collection variables from table 4 having a nonrandom distribution with these same temporal, functional, and spatial divisions. Ordinal-level correlation coefficients between substantive variables in table 3 and collection variables in table 4 that have a significant relationship are shown in table 7.

Such significant temporal, spatial, and functional relationships with substantive variables might be of interest, or might be confounding variables, depending on the research question. To examine change through time in the proportion of hafted tools in the assemblage the strong spatial variability in proportions of hafted tools (table 5) is a confounding variable. If a dependent variable of interest has a significant relationship with a confounding variable, then this relationship must be controlled for in some manner.

As an example of what these tables indicate and how they may be used, we might be interested in the proposition that the degree of craft

Table 5. Relationships between major temporal, spatial, and functional classes and the substantive variables from table 3

	Period	Early/ late	Local- ities	Upstream/ downstream	Habitation/ seasonal
Ceramics					
PSM > 4	.03	.00			
"Late" types	.04	.02			
Modified items	.06	.06			
Jars		.09			
Flaked lithic tools					
Mean weight		.04	.08	.06	.03
Utilized flakes			.06	.04	.08
Nonflaked lithic tools					
Hafted items		.04	.03	.01	
Well shaped or stylized	.04	.02		.02	
Resurfaced with light vertical force				.08	
Resurfaced with medium vertical force		.10			
Resurfaced with heavy vertical force	.06	.01		.02	

NOTES: Entries appear in the matrix for row variables having a significant relationship with column variables at the .1 alpha level or less. The entries in the matrix are the probabilities that such a relationship could have been achieved by chance alone in a sample drawn by independent randon sampling from a population in which no such relationship existed.

The test statistic is a Chi-square approximation of the Kruskal-Wallis one-way nonparametric analysis of variance.

Sample size for most comparisons is 30, with the element data for Modeling Period 7 excluded.

PSM - Production step measure.

Table 6. Relationships between temporal, spatial, and functional classes and the collection variables from table 4

	Modeling period	Early/ late	Local- ities	Upstream/ downstream	Habitation/ seasonal
Study unit type					
Surface structures		.02			
Pitstructures					.05
Nonstructures		.07			
Refuse type					
De facto refuse	.10	.04			
Mode of collection					
Screened	.02	.00	.02	.00	

NOTES: Sample size and test statistic are as in table 5.

Collection variables were computed for the four major material classes. Those shown represent the proportions of ceramics recovered. They are generally representative of the results for flaked lithic tools, debitage, and nonflaked lithic tools as well.

Table 7. Correlations between collection variables and substantive variables

	Surstr	Pitstr	Nonstr	mgs	De facto	Use assoc	Screen	Trowel
Ceramics								
PSM ≥ 4	.25		.27		.31			
Dolores Tract	-.28							
"Late" types	.35		.33					
Modified items	-.30			.32				
Jars					-.28	-.35		
Flaked lithic tools								
Mean weight					.28	.34		
Utilized flakes								-.29
Projectile points							-.24	
Complete or nearly							.25	
Nonflaked lithic tools								
Mean weight		.24		-.32				
Complete or nearly				-.46	.24			.29
Hafted items							.41	
Ornaments		.29						
Hammerstones				.27			.32	
Well shaped	.25				.31			.26
Resurfaced/LVF				-.24				
Resurfaced/MVF		.42						.27
Resurfaced/HVF		.45			.32		.42	
Debitage								
Mean weight	.24		-.26	.36	-.26		-.30	
Cortex				.25				

NOTES: Values shown are Kendall Tau-b coefficients. Only correlations significant at an alpha level of .1 or higher are shown.

More information on the substantive variables can be found in table 3.

Pitstr - Pitstructure.
Surstr - Surface structure.
Nonstr - Nonstructural.
mgs - Modern ground surface.
assoc - Association.
PSM - Production step measure.
LVF - Light vertical force.
MVF - Medium vertical force.
HVF - Heavy vertical force.

specialization in the manufacture of decorated ceramics increased through time during the occupation of the river valley. Perhaps we wish to examine the proportion of ceramics with a PSM (production step measure) of four or more as one implication of this hypothesis. (Although we are examining only one variable, this is an assemblage level question since it deals with the relative frequency of one set of ceramics in the collection in relation to the entire sample of ceramics.) Table 5 shows that, indeed, a nonparametric one-way analysis of variance is able to reject the null hypothesis of no difference in this proportion among periods (alpha = .03) and between the early and late divisions (alpha = .00). Are there any confounding variables that must be considered? The correlation analysis using Kendall's Tau-b (table 7) reveals that this proportion is significantly correlated with the proportion of ceramics representing de facto refuse (alpha = .02) and the proportion of ceramics from nonstructures (alpha = .06) and surface structures (alpha = .10). Is this in itself a matter for concern?

Table 6 shows that each of these three potential confounding variables changes significantly through time as does the variable of interest. Then, is the change in PSM through time due to systemic changes in ceramic manufacturing as specified by our research hypothesis? If PSM is higher late in time because it is always higher in surface structures, nonstructures, and de facto refuse, and because there are relatively greater amounts of de facto refuse, nonstructures, and surface structures in elements dating to late periods, then these are not confounding independent factors but merely intermediate carriers of the effect of change in time on the PSM. (Certainly we suspect this is true of surface structures during this period of pithouse-to-pueblo transition.) If, on the

other hand, the increase in PSM is merely due to sampling effects resulting from collecting a higher proportion of de facto refuse or from conducting more nonstructural excavation in elements that happen to be late, then these are confounding variables that may cause us to mistake a sampling effect for a real change through time.

Two options are available in pursuing this last question. It might be possible to eliminate from consideration elements that were not relatively homogeneous on proportions of materials from nonstructural excavation and from de facto refuse. In this case, however, such elimination would severely reduce the already small sample size. More realistically, changing values of the confounding variables may be controlled statistically to determine whether change through time still has a significant relationship with PSM when variability due to the confounding variables is held constant. These approaches are being considered in a separate paper (Kohler et al. 1983). Further analysis along the lines of this example suggests that change in PSM is best explained by a model incorporating time and the proportions of nonstructural materials as independent variables. However, when the differential effect of nonstructural materials is controlled for, there still remains a significant relationship between time and PSM.

DISCUSSION AND CONCLUSIONS

Despite efforts to standardize the recovery of data in the field and to further select highly interpretable temporal-spatial subsets of sites, not all intercollection variability that results from sampling effects, broadly defined, can be eliminated. Users of these data--and certainly all other archaeological data as well--should remember three guidelines in hypothesis testing:

1. When possible, use variables that appear to be relatively unaffected by collection effects. This includes variables shown in table 3 that do not appear in table 7, such as flaked lithic tools of non-local materials or well-shaped, stylized tools. Even such apparently "safe" variables may, however, have confounding relationships with systemic phenomena not specified in the research hypothesis.
2. When variables that are sensitive to collection or systemic confounding effects must be used, select cases that are as homogeneous as possible on the confounding variables to minimize interference. In the example developed in the previous section, it is clear that if we wished to use PSM to discover "nonegalitarian sociopolitical structures" as done by Upham et al. (1981) then we would not be able to pool our early and late periods in the analysis, and, ideally, periods would be analyzed separately.
3. When neither of the first two guidelines is practical, it may be possible to statistically control for the confounding variables. Possible techniques for this include partial correlation analysis, analysis of covariance, or multivariate categorical analyses, depending on the characteristics of the data under consideration.

We have presented more data here than we have discussed. The purpose of this paper is to alert DAP workers to variables that are especially subject to vagaries of collection techniques and variability across temporal, spatial, and functional boundaries. In some variables and units of analysis examined, as much as 35 percent of the variation seems to be due to collection techniques or related confounding factors rather than to systemic variability. Other variables appear to be relatively immune to such effects. Some aspects of the information presented, which cannot be further discussed here, are relevant to an understanding of site formation processes, or to substantive studies of geographic, temporal, and functional variability within the DAP.

We believe that this study points up the necessity for archaeologists to be suspicious of their data and cautious in its use. We urge others to do what we have done here for their own data, or to go beyond this analysis, as we are trying to do, with more rigorous approaches to explaining variabilty, such as path analysis and other variants of causal analysis.

REFERENCES CITED

Arnold, Dean E.
1978 Ceramic variability, environment and culture history among the Pokom in the valley of Guatemala. In The spatial organization of culture, edited by I. Hodder, pp. 39-59. University of Pittsburgh Press, Pittsburgh.

Binford, Lewis R., and Sally R. Binford
1966 A preliminary analysis of functional variability in the Mousterian of Levallois facies. In Recent studies in paleo-anthropology, edited by J. D. Clark and F. C. Howell. American Anthropologist 68:238-295.

Blalock, Hubert M., Jr.
1979 Social Statistics (revised second ed.). McGraw-Hill, New York.

Blinman, Eric
1981 Dating with neckbands: calibration of temporal variation in neckbanded ceramic types. In Dolores Archaeological Program synthetic report, 1978-1981 DAP-055, pp. 346-374. Final report submitted to the U.S. Bureau of Reclamation, Upper Colo. Region, Salt Lake City, in compliance with Contract No. 8-07-40-S0562.

1982 Ceramic dating of survey collections. Ms. on file, Dolores Archaeological Program, Dolores, Colorado.

Blinman, Eric, and C. Dean Wilson
1982 Operationalizing the ceramic production step measure. Ms. on file, Dolores Archaeological Program, Dolores, Colorado.

Bordes, Francois
1972 A tale of two caves. Harper and Row, New York.

Cook, Sherburne F.
1972 Can pottery residues be used as an index to population? Contributions of the University of California Archaeological Research Facility 14:17-40.

Englebrecht, William
1978 Ceramic patterning between New York Iroquois sites. In The spatial organization of culture, edited by I. Hodder, pp. 141-152. University of Pittsburgh Press, Pittsburgh.

Feinman, Gary M., Steadman Upham, and Kent G. Lightfoot
1981 The production step measure: an ordinal index of labor input in ceramic manufacture. American Antiquity 46:871-884.

Ford, James A.
1952 Measurements of some prehistoric design developments in the southeastern states. American Museum of Natural History Anthropological Papers 44(3).

Kane, Allen E., and Christine K. Robinson
1980 Pollen and bulk soil sampling strategies, Dolores Archaeological Program, southwestern Colorado. Paper presented at the 53rd Annual Pecos Conference, Mesa Verde National Park, Colorado.

Kane, Allen E., Nancy J. Hewitt, Joel M. Brisbin, Patrick Hogan, Gary A. Brown, Steven D. Emslie, J. Holly Hathaway, Frank C. Leonhardy, William A. Lucius, and Roger A. Moore
1981 The Dolores Archaeological Program field manual - 1978. Dolores Archaeological Program Technical Reports DAP-011. Final report submitted to the U.S. Bureau of Reclamation, Upper Colo. Region, Salt Lake City, in compliance with Contract No. 8-07-40-S0562.

Kane, Allen E., William D. Lipe, Ruthann Knudson, Timothy A. Kohler, Steven E. James, Patrick Hogan, and Lynne Sebastian
1981 The Dolores Archaeological Program research design. Dolores Archaeological Program Technical Reports DAP-002. Final report submitted to the U.S. Bureau of Reclamation, Upper Colo. Region, Salt Lake City, in compliance with Contract No. 8-07-40-S0562.

Kohler, Timothy A., and G. Timothy Gross
1981 Probability sampling in excavation: a program review. In Dolores Archaeological Program synthetic report, 1978-1981 DAP-055, pp. 169-181. Final report submitted to the U.S. Bureau of Reclamation, Upper Colo. Region, Salt Lake City, in compliance with Contract No. 8-07-40-S0562.

Kohler, Timothy A., Carl J. Phagan, and Eric Blinman
1983 Sources of confounding variability in archaeological collections. Ms. on file, Dolores Archaeological Program, Dolores, Colorado.

Lewis-Beck, Michael S.
1980 Applied regression: an introduction. Series: Quantitative Applications in the Social Sciences 22.

Longacre, William A.
1970 Archaeology as anthropology: a case study. Anthropological Papers of the University of Arizona 17.

Nance, Jack D.
1981 Statistical fact and archaeological faith: two models in small-sites sampling. Journal of Field Archaeology 8(2):151-165.

Otto, J. S.
1975 Status differences and the archaeological record: a comparison of planter, overseer, and slave sites from Cannon's Point Plantation (1794-1861), St. Simon's Island, Georgia. Unpublished Ph.D. dissertation, Department of Anthropology, University of Florida, Gainesville.

Phagan, Carl J.
1981 Lithic profiles. In Dolores Archaeological Program synthetic report, 1978-1981, DAP-055, pp. 393-412. Final report submitted to the U.S. Bureau of Reclamation, Upper Colo. Region, Salt Lake City, in compliance with Contract No. 8-07-40-S0562.

Plog, Fred T.
1974 The study of prehistoric change. Academic Press, New York.

Rice, Prudence M.
1981 Evolution of specialized pottery production: a trial model. Current Anthropology 22:219-240.

Schiffer, Michael B.
1975 The effects of occupation span on site content. In The Cache River Archeological Project: an experiment in contract archeology, assembled by Michael B. Schiffer and John H. House. Arkansas Archeological Survey, Research Series 8:265-269.

1976 Behavioral archeology. Academic Press, New York.

Schlanger, Sarah, and Patrick L. Harden
1983 1979 and 1980 probability survey of Cline Crest, Grass Mesa, Beaver Point, Trimble Point, and Hoppe Point Localities. Dolores Archaeological Program Technical Reports, DAP-118. First draft submitted to the U.S. Bureau of Reclamation, Upper Colo. Region, Salt Lake City, in compliance with Contract No. 8-07-40-S0562.

Sears, William H.
1973 The sacred and the secular in prehistoric ceramics. In Variation in Anthropology, edited by D. Lathrap and J. Douglas, pp. 31-52. Illinois Archaeological Survey.

Upham, Steadman, Kent G. Lightfoot, and Gary M. Feinman
1981 Explaining socially determined ceramic distributions in the prehistoric plateau Southwest. American Antiquity 46:822-833.