INTRODUCTION

Recently Dunnell (1986:149) identified me as one of those who think that "unit formation is a necessary but intellectually uninteresting activity without major significance for ... (archaeology's) primary goals." I should make my views about classification more clear. I do not find classification very interesting as a topic in its own right, but I agree that it is usually misunderstood and often misused, and I most emphatically think that efforts toward our primary goals (e.g., explanations of cultural change and stability, culture-historical reconstructions that go significantly beyond mere time-space frameworks, believable approximations to past cultural worlds, etc.) are irremediably flawed if they rest on inadequate or misconceived typological foundations. Thus, while classification may not excite all of us as an end in itself, we cannot simply neglect it in the pursuit of more intriguing or more fashionable goals. I would not go quite so far as those who argue that "classification is the most critical and pressing issue in the field" (Dunnell 1986:150), but I agree that it is one of our most critical and pressing issues. If we cannot get our ideas about classification right, we cannot get anything else right.

Archaeologists classify many kinds of entities, including site features, sites, societies, institutions, and cultures. However, the cognitive processes, skills, and many of the issues involved in classifying and sorting artifacts are very different from those brought into play in attempting to define, for example, types of societies. It is a serious mistake to think that "classification is just classification" and to confound concepts and methods applicable to one kind of entity with those applicable to some other kind. In this paper I will concentrate on the classification of artifacts, especially ceramics and lithics. Dunnell (1986) has provided a very useful review of the subject, with a lengthy bibliography. I will not try to duplicate that. I will emphasize certain topics that are especially important because they are in need of clarification, and would repay further study.

1. Why is there such a disparity between the techniques used or prescribed in the literature on formal approaches to classification and the practices actually followed by persons confronted with thousands or millions of objects?

2. There are great differences between looking for structure within an assemblage from a single period at a single site, looking for structure that is shared by a number of assemblages, and making comparisons between different assemblages. Obvious as this point may seem, it has not always been appreciated, and this has led to unnecessary muddles on such issues as whether types are discovered or invented and whether cultural change is thought to be gradual or punctuated by abrupt shifts.

3. It is often of interest to determine the extent to which a collection of objects can be sorted into relatively discrete groups—that is, groups whose members resemble one another decidedly more than they resemble objects in other groups. The degree of such discreteness is a discoverable property of a given collection described in terms of a chosen set of variables. Great confusion has been caused by confounding this kind of discreteness with the intrinsic logical discreteness of nominal (and from certain viewpoints, ordinal) measurement scales. This confusion needs to be cleared up in order to understand what should and should not be meant by the phrase "internal cohesion/external isolation", a concept basic to all further discussion, which I will abbreviate as IC/EI.
4. It has become a truism that different classifications would be better for different purposes, but most publications on artifact classification still tend to concentrate on a particular purpose and give at best perfunctory attention to classifications adapted to other aims. I will discuss the contrasting implications for classification of three purposes that are especially important: the dating of sites or parts of sites with which a collection of objects is associated, the attempt to identify and trenchantly describe consistent practices of ancient artifact makers (which can be a basis for inferences about their own system of classification), and inferences about the transmission of ideas over space and time. For these different purposes, it is not merely that different properties of the objects may be more significant for one purpose than for another; there are also crucial differences in the most useful approaches to classification.

5. Most formal analyses have concentrated on improving ways to define or discover what can be called "minimal" classes; that is, ways to sort objects into groups whose internal variability can be considered unimportant. We need to complement this with more attention to improving ways to handle complex information about the similarities and differences between various minimal classes. This is especially important for comparisons between assemblages intended to elucidate the transmission and modification of ideas across space and time. Resemblances between classes can be handled to some extent by studies of 'modes' (in Rouse's sense) and by using a hierarchical classification, but more powerful concepts and techniques are needed.

SOME PRELIMINARIES

Throughout I follow Rouse (1939:11) and Dunnell (1971), at least approximately, in making a distinction between the formation and definition of classes and the use of these classes in the sorting of objects into groups. Classes are concepts that are defined, groups are sets of objects that can be described. A collection of objects can be sorted into groups according to a set of classes. Broadly, there are two major ways of defining classes. One way is to define them more or less independently of any specific body of material, in terms of properties deemed relevant for some specific interest of the classifier. Examples would be intended use, value or preciousness, or provenance of raw materials. The other way is to sort a specific collection of objects into groups whose members (in terms of properties salient to the analyst) tend to be much like one another and considerably unlike objects assigned to other groups, and then trying to see what distinguishes the groups from one another. Succinct statements of what differentiates the groups from one another can be phrased as class definitions.

I use the term 'property' as an unanalyzed concept, meaning rather vaguely anything we might notice about an object, such as that it is red, or 12.3 cm long, or has a streakily burnished surface. In a well designed descriptive system, properties should be organized in terms of variables and states (or values) of these variables.

Following Spaulding (1982 and earlier) I use the term 'variable' to refer to any distinct type of property of an object, such as color, length, type of surface finish, presence of handles, and so on. Each variable has at least two possible states or values. Thus, the variable 'color' can have such states as 'red', 'brown', 'black', 'white', and so on; while 'length' can have a very large number of states, such as '6.4 cm', '6.5 cm', etc.; and 'surface finish' can have values such as 'matte', 'streakily burnished', 'well-burnished', 'polished', and so on.

States of variables measured on a nominal scale may be called 'attributes'. I emphasize this because some literature (see Hodson 1982; Djindjian, this volume) uses the term 'attribute' to mean a nominal variable, thus introducing a potential for fundamental misunderstandings. I will use the term 'attribute' only to mean a particular value of a nominal variable.
In a properly designed descriptive system, any one object can manifest exactly one state of each variable. An object cannot be both 6.4 and 7.3 cm long. It can, of course, be both red and white, but this simply means that the variable 'color' should be redefined, probably by allowing for various color combinations; or perhaps several variables could be defined, such as 'primary color', 'secondary color', and so on. For all variables the value 'missing data' will be needed in order to deal with damaged objects, 'other' will be needed for many nominal variables, and 'not applicable' may be required as a possibility when one variable, though logically distinct, is in some way contingent on the state of another variable.

Following Adams (1988) I use 'intrinsic variable' to refer to some property of an object such as length or color that inheres in the object itself and is independent of the properties of other objects or the deposit with which the object is associated. Variables that characterize the deposit or the collection that includes the object are 'contextual variables' and those that relate to inferred properties such as date or use of the object are 'inferential variables'. Note that although we cannot characterize collections of objects without taking account of the intrinsic properties of the objects that comprise the collections, many of the properties we are interested in are properties of collections rather than of objects. This is emphatically so for IC/EI, to be discussed below. It cannot be an intrinsic property of single objects; it always depends on the particular set of objects studied jointly in a particular case.

THE DISPARITY BETWEEN CUSTOMARY AND FORMAL APPROACHES

Dunnell (1986:149-150) comments on the apparently low level of interest in artifact classification, and observes that "The 'theoretical' literature has diverged from practice to such a degree that the two are now unrelated". Others, such as Hodson (1982:28) and Aldenderfer (1987:26-27) also note this. Many archaeologists seem not to have thought very hard about classification, and continue to use customary practices, either because they do not really perceive any serious shortcomings in these practices or because, if they do, they have not seen anything in the theoretical literature that looks really helpful. They apparently feel that the recent theoretical literature on artifact classification is too involved with dubious and sometimes poorly understood philosophy, excessively esoteric mathematics, or the constraints of computer technology to be of much use to them in dealing with the practical tasks they confront. In 1988 the Newsletter of the Classification Society of North America, which publishes the Journal of Classification, listed only seven members who identify themselves as anthropologists and two archaeologists, out of a total of about 406 members. Most are psychologists, statisticians, or biologists. This suggests that, although there is a reasonable amount of interest in mathematical methods of classification in North America, very little of this interest comes from archaeologists. Indeed, it implies that even among quantitative archaeologists, very few subscribe to the Journal of Classification. Either we are missing something important, or, as I suspect, few of the articles in that journal are of much help for the archaeological problems of most interest to us at present.

While most archaeologists remain inarticulate on these issues, W.Y. Adams (1988; see also Adams and Adams, in press) has recently published a critique of theoretical trends as he views and offered an unusually thoughtful defense of traditional approaches. I have some disagreements with his discussion. For example, he conjures up an image of a cohesive and highly influential school of formal classifiers who have a misplaced confidence in computer methods that has drawn attention away from sounder and more realistic approaches to classification problems. This is very different from my own perception that those who write about theory of artifact classification have spent much of their time arguing with one another and have had very little impact on practice. Also,
although in the past some expected too much from computer methods, many archaeologists working with quantitative and formal approaches have themselves recently emphasized some limitations of these approaches (e.g., Aldenderfer 1987:26-27; Read 1989).

Although the methods Adams defends seem to work all right for chronology, he does not deal adequately with the issues raised by other archaeological purposes. Nevertheless, his discussion is very stimulating and it prompts a closer look at the strengths as well as the weaknesses of the procedures used by archaeologists in customary, non-formal approaches.

It is worth listing more explicitly some of the shortcomings to which customary procedures are prone.

1. Publications on "the ceramics" or "the lithics" or "the artifacts" that satisfy the standards of peers in fact needlessly throw away a great deal of very useful information. The analyst stops short of investigating some interesting problems for which the materials are relevant, and presents the data in ways that make it impossible or at least exceedingly difficult for anyone else to use them to investigate these problems.

2. Information that does find its way into reports is unnecessarily hard to get at.

3. Sometimes data may actually be misrepresented. For example, modal or typical properties of objects placed into a single group may be reported without adequate description of ranges of variation. Another common form of misrepresentation is what Hill and Evans (1972) call 'reification'. Named types take on lives of their own and are treated as simple units of data rather than complex and provisional interpretations of data. Among other things, these practices make it particularly difficult to ascertain the actual degree of IC/EI represented by particular groups of objects in particular collections.

4. The analyst investigates interesting problems but makes faulty arguments about the implications of his or her data for these problems.

5. Results as good as or better than those customarily obtained could be gotten by more cost-effective procedures. This is a point rarely mentioned, but emphasized by Adams (1988).

The rhetoric of the formal literature has tended to suggest that customary procedures are simplistic, or, at any rate, deficient in logical rigor. For example, Dunnell (1986:178) says "in spite of the demonstrable utility of the culture-historical types, ... selecting definitive combinations of modes or attribute classes was still a matter of trial and error combined with whatever prior knowledge might be relevant in a given case". The suspicion that what is usually done is rather simple-minded is further encouraged by the fact that so many published descriptions of archaeological materials, especially ceramics, either do not deal with the rationale of the classifications at all, or else do so in ways that do not do justice to what the authors have actually accomplished. Discussions of the logic of the procedures are often obscure or muddled, descriptions of the procedures themselves are often excessively vague and sometimes unintentionally misrepresent what was really done, and concepts and techniques that work reasonably well for certain purposes are promoted as general purpose procedures—as "the" way to do classification and sorting. In fact, highly complex cognitive skills are usually involved, and many traditional classifiers are better and more sophisticated in doing what they do than in describing what they do. Artifact classification poses a number of interesting problems in cognition.

Rather than beginning with certain mathematical procedures and asking what uses we can find for them, it would be better to begin with the problems we actually experience in archaeological classification and sorting, and then ask how formal procedures might help. What really happens when one either learns or creates a system of classification? I will not try discuss this in detail here, but I will offer some thoughts derived from attempts to
comprehend my own experiences in learning other peoples' typologies, creating my own, describing objects, and sorting them into groups on the basis of typologies.

A cautious and limited comparison with language learning seems useful. Attempts to make close parallels with language are likely to be unprofitable, leading to false analogies or worse, but nevertheless I strongly suspect that learning or creating artifact classifications draws importantly on some of the innate human skills and capacities that are basic to language behavior. At the same time, introspection tells me that learning to recognize a particular type with high accuracy also has much in common with learning to ride a bicycle. There was a time when I could not, and then, somehow, I could, and since then I have not understood how not to ride a bicycle (which does not mean I might not fall off in sufficiently difficult circumstances), just as I cannot notice speech differences that are phonemic in my idiolect of English (unless the signal to noise ratio is very low). Similarly, there is something that, for lack of any better term, I call 'sensitivity' involved in learning to identify types. An awareness of what to notice seems to arise that is not entirely captured by any verbal description. For example, it does not take a long time working with the central Mexican pottery called 'Thin Orange' to learn to identify it very accurately even when it is quite thick and greenish gray, or just as accurately to sort many sherds into other groups, even though they happen to be thin and orange. Of course, there is much more to a good description of Thin Orange than saying it is usually orange and sometimes quite thin. Nevertheless, I think it would be extremely difficult and perhaps impossible to put into words everything that has enabled me and others to identify this category with high accuracy and high agreement by just handling and looking at them. I also find it is hard to understand why some people can do a much better job than others at sorting unless it involves something important that cannot be, or at least customarily is not, put into words.

In *Life on the Mississippi* Mark Twain (1874:75-77) tells how, as an apprentice pilot on a riverboat, he nearly ran the boat ashore trying to avoid a non-existent reef. The pilot, Mr. Bixby, says to him:

"Now, don't you see the difference? It wasn't anything but a wind reef. The wind does that."

"So I see. But it is exactly like a bluff reef. How am I ever to tell them apart?"

"I can't tell you. It is an instinct. By and by you will just naturally know one from the other, but you never will be able to explain why or how you know them apart."

"It turned out to be true. The face of the water, in time, became a wonderful book—a book that was a dead language to the uneducated passenger, but which told its mind to me without reserve, delivering its most cherished secrets as clearly as if it uttered them with a voice."

One can carry this line of thought too far. Accepted uncritically, it implies both that there is nothing really wrong with archaeological typology as traditionally practiced, and that it depends on ineffable skills that have little in common with the explicit kinds of reasoning that are essential to any scientific process. To dwell only on 'sensitivity' would be to ignore the very real conceptual advances of formal approaches over the past fifty years. Yet, formal analysts have been too ready to overlook these skills altogether, to depreciate them as "trial and error". To clarify our ideas about archaeological classification we cannot leave out logic, but we also cannot leave out hard-to-articulate pattern-recognition abilities. We have to understand these abilities better, and incorporate this understanding in any adequate account of artifact classification.

It may be important that we can acquire great skill in recognizing types accurately at an age when our skills in language acquisition are much less than those of children. This may mean that what we do in sorting artifacts is more accessible to conscious introspec-
tion. To pursue this a little further, my experience is that, although many sherds that differ considerably from one another can be sorted into the same group with no hesitation, nevertheless there are others that are truly difficult to classify. It seems that what is involved is something like the concept of the "prototypical example" which leads to a fuzzy-edged polythetic class that is defined in terms of a focal point in the multidimensional space comprised by the relevant descriptive variables, rather than by a hard-edged hypervolume in that space. However, to my mind, there are in fact focal volumes, ranges of values of variables within which an object is a "good" case of Thin Orange, or whatever, surrounded by a penumbra of dubious or atypical examples. It is not entirely clear why certain variations are still within the "good" range, while others suffice to make an object marginal. Sometimes a particular variation seems typologically unimportant because we observe a considerable range of values on a single object, as when a Thin Orange bowl is mostly orange but has some greenish areas, which implies that the greenishness represents nothing more than uneven access to a predominantly oxidizing atmosphere in firing, or possibly temperature differences. In other cases archaeometric studies or our general knowledge of the material or experience in the activities for which we think the objects were used may suggest that certain differences are not very meaningful, while other contrasts are more likely to reflect different sources of raw materials, different manufacturing techniques, and/or different uses. Contrasts of the latter kind are more apt to make it problematic whether we can comfortably sort the object into the group it somewhat resembles. Also, any impression that the marginal cases, as a tentative group, differ noticeably from "good" cases in their associations, whether the impression is obtained informally or through statistical analyses, will imply that there is indeed good reason to sort them into a distinct group.

Thus, customary intuitive techniques for classifying and sorting objects rely heavily on highly developed human capacities for pattern recognition. These capacities are still poorly understood. Their great strength is that they far exceed anything we yet know how to formalize or express as computer algorithms. Indeed, as long as we cannot say how we do what we do, we cannot tell a computer how to do it, even if the computer could do it in principle. But the great weakness of our capacities is that because they remain poorly understood they can conceal fallacies and mislead us in ways we find very hard to recognize. For example, as the Rorschach test demonstrates, we can all too easily detect and react to "patterns" and even impose meanings that are in no way inherent in the objects of our attention.

The upshot of the discussion in this section is that studies of human cognition may turn out to be very useful for understanding and improving artifact classification. Artificial intelligence work, particularly trying to tell a computer what it needs to know to reproduce our classes, may also lead to improvements in formal methods.

**DIFFERENT KINDS OF COLLECTIONS**

We can only seek for patterning among variables by comparing the artifacts, as described in terms of those variables, in some set of artifacts. Whatever structure we find among the variables is not an intrinsic property of individual objects as length or color are. It is, instead, a property of a collection of artifacts. For example, we can correctly say that, for a particular artifact, its width/length ratio is 0.37. This is an intrinsic property of the artifact, and remains the case whether we consider the artifact in isolation or as a member of some collection of artifacts. In contrast, the statement "The correlation between length and width is 0.87" can only be meaningful with respect to some particular set of artifacts that includes this particular artifact. If the very same artifact is considered as a member of a differently constituted set (for example, by adding or removing some other artifacts) the correlation between length and width may turn out to be quite different.
although in the past some expected too much from computer methods, many archaeologists working with quantitative and formal approaches have themselves recently emphasized some limitations of these approaches (e.g., Aldenderfer 1987:26-27; Read 1989).

Although the methods Adams defends seem to work all right for chronology, he does not deal adequately with the issues raised by other archaeological purposes. Nevertheless, his discussion is very stimulating and it prompts a closer look at the strengths as well as the weaknesses of the procedures used by archaeologists in customary, non-formal approaches.

It is worth listing more explicitly some of the shortcomings to which customary procedures are prone.

1. Publications on “the ceramics” or “the lithics” or “the artifacts” that satisfy the standards of peers in fact needlessly throw away a great deal of very useful information. The analyst stops short of investigating some interesting problems for which the materials are relevant, and presents the data in ways that make it impossible or at least exceedingly difficult for anyone else to use them to investigate these problems.

2. Information that does find its way into reports is unnecessarily hard to get at.

3. Sometimes data may actually be misrepresented. For example, modal or typical properties of objects placed into a single group may be reported without adequate description of ranges of variation. Another common form of misrepresentation is what Hill and Evans (1972) call ‘reification’. Named types take on lives of their own and are treated as simple units of data rather than complex and provisional interpretations of data. Among other things, these practices make it particularly difficult to ascertain the actual degree of IC/EI represented by particular groups of objects in particular collections.

4. The analyst investigates interesting problems but makes faulty arguments about the implications of his or her data for these problems.

5. Results as good as or better than those customarily obtained could be gotten by more cost-effective procedures. This is a point rarely mentioned, but emphasized by Adams (1988).

The rhetoric of the formal literature has tended to suggest that customary procedures are simplistic, or, at any rate, deficient in logical rigor. For example, Dunnell (1986:178) says “in spite of the demonstrable utility of the culture-historical types, ... selecting definitive combinations of modes or attribute classes was still a matter of trial and error combined with whatever prior knowledge might be relevant in a given case”. The suspicion that what is usually done is rather simple-minded is further encouraged by the fact that so many published descriptions of archaeological materials, especially ceramics, either do not deal with the rationale of the classifications at all, or else do so in ways that do not do justice to what the authors have actually accomplished. Discussions of the logic of the procedures are often obscure or muddled, descriptions of the procedures themselves are often excessively vague and sometimes unintentionally misrepresent what was really done, and concepts and techniques that work reasonably well for certain purposes are promoted as general purpose procedures—as “the” way to do classification and sorting. In fact, highly complex cognitive skills are usually involved, and many traditional classifiers are better and more sophisticated in doing what they do than in describing what they do. Artifact classification poses a number of interesting problems in cognition.

Rather than beginning with certain mathematical procedures and asking what uses we can find for them, it would be better to begin with the problems we actually experience in archaeological classification and sorting, and then ask how formal procedures might help. What really happens when one either learns or creates a system of classification? I will not try discuss this in detail here, but I will offer some thoughts derived from attempts to
comprehend my own experiences in learning other peoples' typologies, creating my own, describing objects, and sorting them into groups on the basis of typologies.

A cautious and limited comparison with language learning seems useful. Attempts to make close parallels with language are likely to be unprofitable, leading to false analogies or worse, but nevertheless I strongly suspect that learning or creating artifact classifications draws importantly on some of the innate human skills and capacities that are basic to language behavior. At the same time, introspection tells me that learning to recognize a particular type with high accuracy also has much in common with learning to ride a bicycle. There was a time when I could not, and then, somehow, I could, and since then I have not understood how not to ride a bicycle (which does not mean I might not fall off in sufficiently difficult circumstances), just as I cannot not notice speech differences that are phonemic in my idiolect of English (unless the signal to noise ratio is very low). Similarly, there is something that, for lack of any better term, I call 'sensitivity' involved in learning to identify types. An awareness of what to notice seems to arise that is not entirely captured by any verbal description. For example, it does not take a long time working with the central Mexican pottery called 'Thin Orange' to learn to identify it very accurately even when it is quite thick and greenish gray, or just as accurately to sort many sherd into other groups, even though they happen to be thin and orange. Of course, there is much more to a good description of Thin Orange than saying it is usually orange and sometimes quite thin. Nevertheless, I think it would be extremely difficult and perhaps impossible to put into words everything that has enabled me and others to identify this category with high accuracy and high agreement by just handling and looking at them. I also find it is hard to understand why some people can do a much better job than others at sorting unless it involves something important that cannot be, or at least customarily is not, put into words.

In *Life on the Mississippi* Mark Twain (1874:75-77) tells how, as an apprentice pilot on a riverboat, he nearly ran the boat ashore trying to avoid a non-existent reef. The pilot, Mr. Bixby, says to him:

> "Now, don't you see the difference? It wasn't anything but a wind reef. The wind does that."
> "So I see. But it is exactly like a bluff reef. How am I ever to tell them apart?"
> "I can't tell you. It is an instinct. By and by you will just naturally know one from the other, but you never will be able to explain why or how you know them apart."
> "It turned out to be true. The face of the water, in time, became a wonderful book—a book that was a dead language to the uneducated passenger, but which told its mind to me without reserve, delivering its most cherished secrets as clearly as if it uttered them with a voice."

One can carry this line of thought too far. Accepted uncritically, it implies both that there is nothing really wrong with archaeological typology as traditionally practiced, and that it depends on ineffable skills that have little in common with the explicit kinds of reasoning that are essential to any scientific process. To dwell only on 'sensitivity' would be to ignore the very real conceptual advances of formal approaches over the past fifty years. Yet, formal analysts have been too ready to overlook these skills altogether, to depreciate them as "trial and error". To clarify our ideas about archaeological classification we cannot leave out logic, but we also cannot leave out hard-to-articulate pattern-recognition abilities. We have to understand these abilities better, and incorporate this understanding in any adequate account of artifact classification.

It may be important that we can acquire great skill in recognizing types accurately at an age when our skills in language acquisition are much less than those of children. This may mean that what we do in sorting artifacts is more accessible to conscious introspec-
Thus, associations between variables depend on choices about relevant data sets, and are not intrinsic properties of single artifacts.

The significance of this point is often not heeded. Writers pay little attention to the kind of data set they have in mind, often fail to notice that others have different kinds in mind, and often make unperceived transitions from one kind of data set to another kind. However, the outcome of a search for good classes or for patterned relationships among variables will depend strongly on the kind of data set (or sets) studied. Furthermore, different archaeological problems presuppose or demand different kinds of data sets. Especially important is the contrast between a data set pertaining to a single assemblage and a data set representing a number of different assemblages.

One particularly important kind of collection is what I call an *assemblage*. This term has been used with surprisingly diverse meanings. What I mean is a collection that approximates as closely as feasible a sample of the objects used by a community or other small social unit during an interval in which little sociocultural change occurred. This is an ideal that can rarely, if ever, be fully realized, but it is often feasible to achieve quite useful approximations.

An assemblage, as I have defined it, is a key kind of collection because it consists, insofar as possible, of objects in use by people who shared similar techniques and standards or who, if they did not, limited their sharing for social reasons rather than because of distance in space or time. It is related to, but not the same as, the set of objects collected from the surface of a site or excavated from a particular layer. The latter kinds of collections may contain mixtures of objects from several different periods; that is, elements of different assemblages. Also, several distinguishable layers sometimes represent a very brief time span and slight culture change, and the collections from these layers can reasonably be grouped into a single assemblage.

Many debates about classification have been misguided because some persons have been talking about searching for patterning within a single assemblage or among a group of very closely related assemblages while others have had in mind comparisons between or across a number of different assemblages. Within-assemblage studies are analogous to the detailed ethnographic study of a single community. Albert Spaulding (e.g., 1953, 1954, 1960, 1976, 1977, 1982) has especially concentrated on within-assemblage investigations.

Far more often, classification systems are intended to be applicable to many assemblages that resemble one another fairly strongly, but nevertheless differ noticeably from one another. This is particularly true for regional surveys, where individual assemblages are often represented by collections much too small to make elaborate within-assemblage analyses profitable, but it is also often the case for classifications of material from excavated sites. Phrases such as “time-space significance” presuppose classes that are meaningful for a number of assemblages since, by my definition of the term, a single assemblage pertains to a unit within which the extent of space and the duration of time are unimportant.

I have been puzzled and irritated for years by the resistance of many archaeologists to recognizing that issues such as whether types are discovered or imposed, and the possibility of making worthwhile inferences about categories meaningful to the ancient makers and users (“emic” categories) cannot be discussed coherently unless one first makes it clear whether one is talking about classes generated by considering objects from a single assemblage (or at most a group of very closely related assemblages) or classes generated by considering objects from diverse assemblages. Spaulding, for example, has emphasized the former, while Ford (e.g., 1954) had the latter in mind. I suspect that most archaeologists are so deeply committed to using classification for between-assemblage studies that they simply cannot assimilate the idea of classifications dedicated to the search for patterning within a single assemblage.

The rationale for treating groups of artisans in contact with one another as key analytic units at present rests largely on “common-sense” assumptions. These assumptions
need more ethnographic testing and, especially, elaboration. Nevertheless, I am confident that they are correct in their broad outlines. Briefly, within a single community, over a relatively short time, makers and users of artifacts have the opportunity to be aware of differences. They may feel that variation in some properties is unimportant, but, with regard to properties they think are important, they will be conscious of similarities and interested in differences. One kind of interest is in maintaining standards. If one’s society shares an idea of what a certain kind of object should be like, there are reasons to avoid objects that differ too much from that idea, as Rouse (1939) pointed out. By ways that may be subtle or not subtle, deviations from shared standards are likely to be discouraged.

At the same time, it is unsatisfactory for the makers/users if different standards are not noticeably different. For example, it is unlikely that distinct concepts of “small” and “large” projectile points will exist unless the majority of points that are produced with the intention of making them “small” are, in fact, noticeably smaller than those intended to be “large”. That is, if the “small/large” contrast was important to the ancient artisans, it is unlikely that they will have made many points that they, themselves, could not easily sort as “small” or “large”. The only way that I can imagine to make such sorting feasible is to avoid making many points of intermediate size. Such an avoidance is detectable by us today, provided that, in tabulating relative frequencies of different point lengths, we take care to study a collection that reasonably represents the objects produced by a set of artisans who shared fairly similar standards; that is, an assemblage, as Spaulding has repeatedly argued. It is this kind of reasoning that provides the warrant for claims that we can make at least some valid inferences about the emics of the ancient artisans. I repeat that much more ethnographic testing and elaboration of these views is needed, but the required ethnography is quite feasible.

It is a different matter that sometimes distinctions between kinds of objects or whole systems of classes are important to maintain, as ethnic markers (Wobst 1977; Hodder 1982), or for other reasons, including the sheer value of novelty.

The joint effect of all these considerations tends in the direction of producing objects that are relatively similar to certain other objects but also relatively different from all other objects; that is, in the direction of producing an assemblage that can be sorted into groups showing relatively high IC/EI. I am not arguing that all assemblages can be sorted into groups that show high IC/EI. On the contrary, I assume that the people in some communities exhibit these tendencies to different degrees than do people in other communities, and that this will lead to interesting differences in degrees and kinds of IC/EI observed in the resulting archaeological assemblages.

More important for the present discussion is that persons not in direct contact with one another, whose contacts are mediated by tenuous chains of multiple third parties, have little opportunity to know whether they are conforming to the same standards. Also, one artisan is under no pressure to avoid objects not easily sortable as either “this” or “that” by another person with whom she/he is not in significant contact. Under these circumstances, “drift” is possible; objects may be made at one time and place according to rules that differ only slightly from those followed at another time and place.

Thus, a search for groups with high IC/EI in data sets that pool different assemblages over time and/or space is very likely to find only a very low degree of IC/EI, which is tantamount to being unable to define classes with sharply defined borders—that is, to be able only to define classes that are rather arbitrarily imposed. By taking the unit of analysis as something that includes more than extremely closely related assemblages and trying to isolate groups with high IC/EI, one is, so to speak, trying to pin down a moving target. Types that may show high IC/EI and be very “real” in the context of a single assemblage can seem unreal and arbitrary when one tries to use them also to characterize a number of assemblages that cover an appreciable amount of time and/or space. I elaborate on this in the section below on transmission of ideas in time and space.
Besides the fact that studies intended to tease out categories salient to the ancient makers and users must work with collections that approximate assemblages, there is at least one other contrast between within-assemblage and between-assemblage studies that bears on appropriate techniques. For a single assemblage, the opportunity existed for the artisans to have made all objects of some broad category, such as chipped stone or ceramics, virtually alike, and our interest focuses on identifying meaningful differences, which may reflect different intended uses, carry some kind of symbolic meaning, or both. However, the more separated two collections are in space and/or time, the more natural it seems that they should differ, and the more our interest shifts to identifying meaningful resemblances. Typically, we ask whether observed resemblances could easily be explained by constraints of raw materials or intended uses, or as merely accidental, or whether they can be interpreted as good evidence for transmission of ideas across space and time. Techniques well adapted to identifying meaningful differences are not likely to be well adapted to identifying resemblances, and vice versa. In the former, for example, ICtEI plays a large role, but in seeking resemblances among diverse assemblages database searching and data retrieval techniques are more important.

**INTERNAL COHESION/EXTERNAL ISOLATION**

Often a search for types is a search for class definitions such that if objects in the relevant collection or set of very closely related collections are sorted in terms of these classes, the result will be groups in which the objects in each group are all decidedly more similar to one another than they are to any objects in other groups. This concept of 'type' is expressed by the phrase "internal cohesion and external isolation". ICtEI is a property of a specific collection of objects, the objects being characterized by some appropriate set of intrinsic variables. More exactly, it is a property of each of the various groups of objects that can be formed within the collection, since some groups in the collection may exhibit high ICtEI with respect to one another, while other groups in the same collection may exhibit low ICtEI with respect to one another. ICtEI is not an "all or nothing" property; it can range from absent to complete, with many intermediate degrees.

The point to examining a collection with regard to ICtEI is to find out the extent to which there are logically and mechanically possible kinds of objects that are in fact absent or disproportionately scarce in the collection. For example, if a collection has many red vessels and many black vessels, many bowls and many jars, do we find that each color is represented by just about the same proportion of bowls and jars, or do we find that most bowls are black and most jars are red? In the former case, many objects will differ from one another on just one of these two variables and ICtEI will be low, but in the latter case most pairs of objects will either be alike in both color and shape or else different on both variables and ICtEI will be high.

There are good reasons other than discovery of high ICtEI for an archaeologist to form classes of artifacts. Different states of a single variable may be important to distinguish for reasons other than the properties of the collection under study. For example, differences in raw materials can point to different sources or imply different mechanical properties, and thus can be an important sorting criterion if source or mechanical properties are relevant for one's purposes, whether or not groups formed on the basis of raw material show high ICtEI in terms of any other variables.

Discussions of ICtEI have been obscured by not distinguishing between the formal properties inherent in certain scales and the empirically observed frequency distributions in a specific collection of objects measured in terms of these scales. Some types of scales are intrinsically discrete; notably nominal scales in which each value is simply "different" from every other value. Both Hodson (1982) and Spaulding (1982) seem to accept that it can be meaningful to say that a group of objects characterized by a particular set of
values of a set of nominal variables can show high IC/EI simply by virtue of the discrete character of the scales involved.

It would be better to limit the concept of IC/EI strictly to properties of collections, as distinct from properties of scales. There is no way that the frequency distribution of scores of a single nominal variable can show 'external isolation' because there can be no gaps or valleys between peaks. This, in turn, is because the sequence in which values of a nominal variable are listed is arbitrary and therefore no value can meaningfully be "between" any other two values. Objects compared in terms of a single nominal variable can only be scored as 'same' or 'different'; there is no logically possible intermediate score. On the other hand, comparisons of objects in terms of two or more nominal variables makes intermediate degrees of resemblance logically possible. For two nominal variables, for example, a pair of objects can score 'same on both', 'same on just one', or 'same on neither'. In general, for \( M \) nominal variables considered jointly and all weighted equally, there are \( M + 1 \) possible degrees of similarity between a pair of objects. For this reason, the extent to which classes based on two or more nominal variables lead to groups with high IC/EI in a particular collection is an empirical question, depending on the variables and states of variables chosen, the choice of what to include in the collection, and on the properties of the objects included (Cowgill 1982).

In contrast, even a single interval or ratio scale variable can be investigated for the degree of IC/EI the objects in a collection show, by finding out whether there are two or more well defined frequency peaks, separated by ranges of values for which frequencies are much lower.

The issue of nominal variables is somewhat confused by the fact that the decision to characterize a particular property by means of a nominal variable is often based on informally recognizing that if the property had been expressed in terms of an interval variable, or a combination of interval variables, several well-defined frequency peaks would have been observed. For example, it may be quite obvious that a particular collection can be sorted into three shape groups, "plates", "bowls", and "jars", with no intermediate forms. One might, nevertheless, treat shape by means of some rather complex set of interval variables representing diameters at various reference points, ratio of height to some critical diameter, and degree of curvature at various points in the profile. Something as complex as this is in fact done when the purpose is, for example, to seek for distinctions within a particular shape class such as beakers. However, if it is simply obvious that a collection contains no objects intermediate between jars and plates and bowls it may make good sense to shortcut the process and treat the distinct shapes as different states of one simple nominal variable. If the discreteness in terms of interval variables really is so obvious that no elaborate analysis is needed to demonstrate it, it would be pedantic to carry out a formal demonstration.

Nevertheless, there are two warnings to heed. The first is that we must not confound the formal discreetness of all nominal scales with the fact that the choice to express a property by means of a nominal variable rests, reasonably often, on informal observation that the property in question, if it were expressed by means of interval variables, would, empirically, have a highly discontinuous distribution in the collection in question. Inarticulate recognition of this fact probably has much to do with the notion that formal discreetness of a nominal scale has something to do with discreetness of the properties of particular sets of objects.

The other warning comes from the fact that, reasonably often, a nominal variable is used to express a property that, in terms of interval variables, is not clearly discrete in the collection under study. For example, in some collections there may be no intermediate forms between bowls and plates, but in other collections there are. Some of the bowls may be so shallow that they grade insensibly into rather deep plates. If this is the case, it is a fundamental and very serious mistake to begin by treating the plate/bowl contrast as two states of a nominal variable. If the lack of intermediate forms is not so obvious as to
make formal analysis superfluous, one should treat the height-to-width ratio of vessels with
a simple concave profile as an interval variable, plot frequencies, see if more than one
well-defined frequency peak is observed in the collection in question, and decide on this
basis whether and how to replace the original interval variable with states of a derived
nominal variable.

Notice, incidentally, that several kinds of comparisons can be made. Objects can
compared with other objects in the same collection. Objects can be compared with objects
in other collections. Collections, rather than objects, can be considered “cases” and
comparisons made between collections. Classes can be compared with classes, and groups
with groups, either within or between collections.

RESEARCH PURPOSES AND CLASSIFICATION

Convenience in description

It is impossible to handle any sizable collection of archaeological objects without
noticing that some are quite a bit like one another and others are quite a bit different.
Absurd repetition can be avoided and descriptions can be made more orderly and usable if
a general description of each group is given, rather than an object-by-object description.
This remains an important purpose, but compact description should be a byproduct of
classifications intended to be useful for other purposes.

Chronology

We often try to infer the relative frequency with which objects of various types were
used and broken over time, within a restricted region, in order to infer the approximate
date of an assemblage or of a collection associated with a specific layer or site feature,
which in turn can be used to infer a date for the site or layer or feature. The task is to
arrange a number of collections into a convincing time sequence. To do this, one needs
some independent source of at least relative dates for a few of the collections. It usually is
not feasible to get such dates for all the collections, so many must be fitted into the
framework of dated collections solely by means of comparisons between the collections.
This is often done by some form of seriation, in which insofar as possible each collection
is put between the two others it most resembles. Alternatively, one may attempt to define
types and arrange collections so that the percentage distribution for each type has a
compact unimodal “battleship” shape. If similarity between collections is measured by the
Brainerd-Robinson coefficient (Robinson 1951), both approaches lead to similar results.
One must try to avoid confounding differences associated with temporal separation and
those associated with spatial distance or differences in activities. Techniques such as
multidimensional scaling can investigate whether any linear arrangement captures most of
the relationships between the collections, or whether an arrangement in two or more
dimensions is required (Cowgill 1972).

Seriation presupposes arranging a number of assemblages, and IC/EI within a single
assemblage is not very important for identifying the categories most useful for seriation.
Sometimes a single state of a single variable will turn out to have great chronological
value. More often, some combination of two or more variables will have greater signifi-
cance for chronology. Suppose that blackware was popular between AD 500 and 800,
while incised scrolls were popular between AD 700 and 1000. If an assemblage has a high
percentage of both blackware and incised scrolls we can infer that it probably dates
between AD 700 and 800.
One may consider either joint occurrence of specific states of the variables within assemblages, as in the previous example, or joint occurrence on specific objects. The latter lends itself to definition of chronologically significant classes. If the classes also show high IC/EI in most assemblages it will be easier to sort objects, but high IC/EI within assemblages is not essential for a class to have chronological significance.

From the viewpoint of chronology, the most useful variables may be those that do not correlate highly with one another across the set of assemblages under study. In the above example, if we have assemblages spanning the time from well before AD 500 to considerably after 1000, there will be many assemblages with lots of blackware and few scrolls (ca. AD 500 to 700), many with lots of scrolls and little blackware (ca. 800 to 1000), and many with little of either attribute (before 500 or after 1000). The correlation coefficient between percentage of blackware and percentage of incised scrolls across assemblages will be low and negative. Variables will have strong positive correlations across assemblages only if the popularity curves of their various states are nearly synchronous, which means that they provide nearly redundant chronological information.

Behavioral uniformities of artifact makers

It is rarely possible for persons without modern machines to make any two objects that are not noticeably different. Nevertheless, people may make many objects according to the same set of conscious or unconscious procedures, and may perhaps also hold similar "mental templates", so that differences in the products are accidental and neither intentional nor consistent. Some archaeologists, notably Spaulding (1953, 1977, 1982) and Read (1982, 1989) have worked on distinguishing differences in artifacts that reflect differences in consistent practices or templates from those that are accidental. These differences may or may not have been given formal linguistic recognition by the ancient users and makers; all that is usually claimed is that at least they had fairly high salience.

The natural kind of collection for such investigations is a single assemblage, as I explained earlier. The basic idea is that within a community there will have been a limited number of distinct practices and/or templates. The assumption is that differences due to differences in practices or templates will be large relative to accidental differences in following any one practice or template. If this is actually the case, one should be able to define classes in a way that will be enable one to sort the assemblage into groups with high IC/EI.

Spaulding (1982 and personal communication) argues that in the search for behavioral uniformities, in the case of interval variables, one should begin by looking for well-defined multiple peaks in the frequency distribution of scores of each single variable. If a variable has only one well-defined peak it should be temporarily set aside, although it can be examined for possible joint multipeakedness with other variables at a later stage of the investigation. Spaulding argues that if an interval variable shows two or more well defined frequency peaks, it should be re-expressed as a nominal variable with a different state for each peak of the interval variable. This means that the search for patterning in the collection becomes the multivariate investigation of nominal variables only. A number of techniques of discrete multivariate analysis (DMA) are applicable, such as loglinear methods and hierarchical models (Lewis 1986, Spaulding 1982). Spaulding adds that there are other useful levels of analysis for which interval variables are best used in their original form, rather than being used to derive nominal variables.

If an assemblage can be sorted into groups that show strong IC/EI, then the variables used for the sorting, at least if they are nominal, must necessarily show strong relationships when subjected to DMA (Cowgill 1982). However, DMA can also show strong relationship that cannot be expressed in terms of groups with strong IC/EI and that would not be found by object-clustering approaches. For example, a group of objects character-
ized by attributes that are rare in the total assemblage can be uncommon and show little or no external isolation, yet manifest a tendency to favor a particular combination of attributes far more than would be expected if the values of the variables were irrelevant for one another. If, of 200 objects, twenty are red and twenty are bowls, we would expect only two red bowls if there were no relation between shape and color. Finding eight red bowls is strong evidence that makers were less reluctant to use red on bowls than on jars, even though the same collection includes 12 black bowls and 12 red jars besides 168 black jars, and even though it would not be very sensible to define 'red bowl' as a class, since it does not enable one to sort objects, in this collection, into a group with high IC/EI.

Read (1989) points out that the search for behavioral uniformities tends to lead to “taxonomic” classifications, in which the most relevant variable at a given stage of class definition depends on the particular combination of values of variables used to define the more inclusive class that is being subdivided at the present stage. A logical further step would be a “design-grammar” approach (Chippindale 1986) in which emphasis moves away from classification and toward formulating rules that will generate the observed patterning in an assemblage.

If we find that we can indeed formulate rules that go far toward generating the observed pattern in an assemblage, what can we infer about the extent to which the makers of the artifacts consciously recognized these rules? This is a difficult question that requires much ethnographic research.

Transmission of ideas in time and space

Attempts to infer transmission of ideas must involve comparisons between assemblages, in contrast to efforts to infer local customs, which are best carried out by within-assemblage investigations. Sometimes extremely remote connections are sought, but usually comparisons are made among some regional set of assemblages within which some degree of idea-sharing can be taken for granted, and problems concern the timing and intermediate links of various connections, and their social purposes and meanings.

Rouse (1939) and Krieger (1944) put great emphasis on “cultural meaning” for types intended to be useful for studying transmission of ideas. Presumably this means that the time-space spread of a type really should reflect transmission of ideas rather than independent inventions and accidental resemblances. The criterion of “cultural meaning” seems to be that the time-space volume occupied by the type should be connected and reasonably compact.

When intuitive archaeological classifications, such as type-variety approaches, go beyond pure chronology they tend to be concerned with evidence about cultural connections, that is, with between-assemblage studies, while formal approaches have been more concerned with inferring patterns of behavior within single, or very closely related, assemblages. Questions of cultural transmission could also profit from formal analysis, but somewhat different approaches will be needed. Formal analysis has emphasized ways to define classes so that one can get groups with high IC/EI. The goal is to get groups of objects which, although no two of the objects are exactly alike, do not differ from others in the same group in any way that can be considered meaningful. This is a legitimate purpose, but it lends itself to representations that ignore differences between objects in the same group and also ignore resemblances between objects in different groups. However, when we compare different assemblages, we normally find groups that have potentially meaningful differences and also potentially meaningful resemblances.

There are two customary ways to deal with this situation. One is to look at modes in the sense of Rouse (1939), important single attributes shared by a number of otherwise diverse groups in different assemblages. The other is to define a nested hierarchy of
classes. Broad widely inclusive classes, such as “wares”, are defined by a few variables, and each is subdivided into subclasses and perhaps subsubclasses, such as “types” and “varieties”, on the basis of additional variables. The idea is that groups that have important differences and also important similarities will belong to different classes at the lowest level, but will belong to the same class at some higher level of the hierarchy.

Both of these devices are useful but have limits. Modal analysis cannot easily deal with more than one variable at a time. Nested hierarchies are limited because a lower level class can only belong to one higher level class, and this framework cannot represent the merging as well as branching of diverse cultural streams. The appropriate biological analogy is not with taxonomic levels between which gene flow is negligible, but rather with interbreeding varieties of a single species. To be sure, one can establish multiple hierarchies that emphasize different aspects of similarity and differences, but these are cumbersome. Adequate representation of situations of intricate branching and merging is a major challenge to those who study the transmission of ideas along complex networks across time and space.

Dunnell (1986:169-170, 171, 177, 181, 182) seems to argue that the search for IC/EI, or at any rate the belief that IC/EI can be interpreted as reflecting customs or concepts of ancient artisans, has to be associated with an “essentialist” ontological position, in which variation within groups is regarded as nothing but meaningless “noise”. This, in turn, he argues, makes it impossible to think about change; that is, about how the situation represented by one assemblage could have been transformed into that represented by some other assemblage—two assemblages can only be regarded as different. He contrasts this with the view labelled “materialist” in biological and philosophical literature (an unfortunate term, since it seems to have little to do with materialism in the sense of Marx or Harris), which focuses on variation as a critical basis for change.

In fact, one can think it important to investigate assemblages, in terms of relevant variables, in order to ascertain the degree to which strong IC/EI is found, without being an essentialist. One interesting property of an assemblage is the nature and composition of artifact groups within it generated by IC/EI-seeking techniques. The fact that the resulting description is static does not incapacitate us for dealing with change any more than does the fact that a description of a single artifact is static.

Suppose that the lengths of spearpoints in assemblages A and B have frequency distributions as shown in Table 1. For assemblage A, an IC/EI-seeking approach would lead to two types; “short” ones 7 to 13 cm long, and “long” ones 14 to 20 cm long. For assemblage B, the same approach would lead to two types, one 8 to 15 cm long and one 16 to 21 cm long.

If we compare these two assemblages in terms of types we are not forced into an essentialist position. That is, we are not unable to talk about how the patterning represented by A might have changed to that represented by B. It is true that no single set of class definitions will enable us to sort both assemblages into groups that are optimal from the viewpoint of IC/EI. For example, the search for high IC/EI in assemblage A would lead us to define the “short” class as spearpoints 7 to 13 cm long, and the “long” class as those between 14 and 20 cm in length, but these definitions do violence to the “natural cleavage plane” in assemblage B, which runs between 15 and 16 cm. They do violence to what it seems felicitous to call “local rules” of manufacture. It was the insistence that one had to apply the same class definitions to all assemblages being compared—a method that works quite well for seriation but is simply inappropriate for many other problems—that led Ford (1954 and elsewhere) to claim that types were entities imposed by the analyst rather than empirical discoveries.

A natural (natural to me, at any rate) and succinct way to express the comparison between spearpoints in these assemblages is to say two distinct things. First, from the viewpoint of IC/EI, in both assemblages spearpoints can be nicely subdivided into “long” and “short” groups, but both groups tend to be somewhat longer in assemblage B. One
could summarize this quantitatively in a number of ways. For example, one could say that short points in A are mostly in the range from 8 to 11 cm with a median length of 9-10 cm, while short points in B are mostly in the range from 9 to 15 cm with a median length of 12-13 cm. Similarly, in A long points are mostly in the 15 to 19 cm range with a median length of 16-17 cm, while in B long points are mostly between 16 and 20 cm with a median length of 18-19 cm. Thus, short points in assemblage B tend to be about three cm longer than in A, while long points tend to be about two cm longer than in A.

The second thing it is important to say, by way of succinct comparative summary, is that about 45% of the points in assemblage A are short, while in B 62% are short. Thus, although short points tend to be somewhat longer in B, there is a higher proportion of them.

It is easy to imagine assemblages where the situation is more complex than in this example even if, as here, the comparison is restricted to a single variable. Joint considerations of two or more variables can also complicate the situation. This example is only a sketch to show that it is quite possible to compare assemblages in a way that (a) does justice to "local rules" by forming classes within each assemblage on the basis of IC/EL, and (b) allows us to describe changes in one assemblage that would lead to the other assemblage, thus freeing us from an essentialist straightjacket.

My example is inductive, empirical, and descriptive, since I have not even hinted at explanations grounded in theory for the change from A to B. The approach is addressed to a "what" question rather than a "how" or "why" question. However, I argue, it is an approach to description that lends itself extremely well to the formulation of appropriate "how" and "why" questions concerning cultural change and cultural interactions. Perhaps a better way to put it is that my example, if it has achieved its purpose, will have vividly exposed inadequacies in trying to make trenchant descriptions of the differences between assemblages if one insists on using identical class definitions for all the assemblages being compared.

Notice that the method I have sketched makes the question of gradual versus abrupt change open to empirical investigation. One should seek for other assemblages that
represent people who are plausibly situated between A and B in networks of human interaction (which usually means intermediate in space/time) and observe the extent to which configurations intermediate between A and B can in fact be discovered. Are we in fact only able to find assemblages highly similar to A and others highly similar to B, or are many finely graded in-between steps observed? If we plot properties of assemblages in space and/or time, do we find gentle gradients or abrupt discontinuities? If we find discontinuities, where and when are they located, and what do they imply about socio-cultural phenomena? Of course, to complicate matters further, we may well find a smooth gradient for one variable or set of variables (e.g. style) concurrently with abrupt change for other variables (e.g. raw materials).

*Other purposes of classification*

Classifications may also be intended to apply to studies of intended uses of artifacts, their actual (primary and secondary) uses, techniques of manufacture, sources of raw materials, styles, and symbol systems. I am not aware of any formal issues that these purposes raise that I have not already touched on. For many of these topics what is most important is learning what to look for and the meaning of what one sees, and one discovers this mainly from ethnography and experimental archaeology. Formal analysis can help, but only if the right observations have been made in the first place.

*CONCLUSION*

In this paper I have not attempted to cover every possible topic concerning artifact classification or to cite every relevant recent publication. I have only sketched those aspects of those topics that I personally find of greatest interest or feel best qualified to discuss. I hope I have chosen well enough so that what I have written will be useful, and perhaps sometimes provocative, to others.

*ACKNOWLEDGEMENTS*

I thank William Y. Adams, Martin Biskowski, Robert C. Dunnell, Albert C. Spaulding, Albertus Voorrips and Robert Whallon for their comments on one or more earlier drafts. I regret very much that ill health prevented Spaulding from commenting on the latest version. I particularly thank Adams for allowing me to make use of his sections of the 1987 draft of Adams and Adams (in press). I have made many changes in response to all these inputs, but responsibility for the final version is, of course, solely mine. I also thank the National Endowment for the Humanities for a Fellowship for University Teachers which aided some stages in the protracted effort of completing this paper.

*NOTES*

1. The degree of discreteness discovered will, of course, differ if one chooses a different set of variables in terms of which to characterize the collection.

2. Giddens (e.g., 1979:5) makes a useful distinction between practical consciousness and discursive consciousness. In those terms, my point is that formal approaches to classification have emphasized that which can be expressed in discourse and have not yet dealt adequately with the role of practical consciousness in archeological practice.
REFERENCES

Adams, W.Y. 1988

Adams, W.Y., and E.W. Adams in press

Aldenderfer, M.S. 1987

Chippindale, C. (ed.) 1986

Cowgill, G.L. 1972

Cowgill, G.L. 1982

Dunnell, R.C. 1971
Systematics in Prehistory. The Free Press, Riverside, New Jersey.

Dunnell, R.C. 1986

Ford, J.A. 1954

Giddens, A. 1979


Hodder, I. 1982

Hodson, F.R. 1982

Krieger, A.D. 1944

Lewis, R.B. 1986

Read, D.W. 1982

Read, D.W. 1989

Robinson, W.S. 1951

Rouse, I. 1939
Spaulding, A.C. 1953

Spaulding, A.C. 1954

Spaulding, A.C. 1960

Spaulding, A.C. 1976

Spaulding, A.C. 1977

Spaulding, A.C. 1982

Twain, M. 1974

Wobst, H.M. 1977